



## **The Ethics of Synthetic Biology Respecting Life and Managing Risk**

Christiansen, Andreas

*Publication date:*  
2016

*Document version*  
Other version

*Document license:*  
[CC BY-NC-ND](#)

*Citation for published version (APA):*  
Christiansen, A. (2016). *The Ethics of Synthetic Biology: Respecting Life and Managing Risk*. Det Humanistiske Fakultet, Københavns Universitet.



# PhD Thesis

Andreas Christiansen

## The Ethics of Synthetic Biology

Respecting Life and Managing Risk

Advisor: Sune Holm

Submitted: 31/08/2016

The Ethics of Synthetic Biology: Respecting Life and Managing Risk

PhD Thesis

By Andreas Christiansen

Department of Media, Cognition and Communication

Section of Philosophy

Academic advisor: Sune Holm

Submitted 31 August 2016

Word count: 50.700

## Table of contents

Acknowledgements	2
Article overview	4
Introduction	6
Article 1 Synthetic Biology and the Moral Significance of Artificial Life	70
Article 2 Similarity Arguments in the Genetic Modification Debate	92
Article 3 Rationality, Thresholds and the Precautionary Principle	126
Article 4 On the Cognitive Argument for Cost-Benefit Analysis	156
Resume	192
Summary	193
References	194

## Acknowledgements

My biggest debt is to my advisor, Sune Holm. Sune has provided detailed comments on the entire thesis and on several drafts that never made it into the final products, which no doubt improved the final product. He has been enormously encouraging and helpful, which has made the transition from being a student to a researcher much easier and less frightening than it might otherwise have been.

Apart from Sune, several people have read parts of the dissertation or earlier drafts. I want to thank Martin Marchman Andersen, Karin Jønch Clausen, Klemens Kappel, Esben Høgh, Tom Douglas, Julian Savulescu, Michael Plant, Carissa Véliz, Norbert Paulo, Maria Serban, Tanja Rechnitzer, Josefine Pallavicini and Johanna Privitera for providing helpful comments and criticisms. I also thank participants in numerous workshops, colloquia and conferences for comments on oral presentations of material included here.

I owe a special thanks to a number of people who have provided support, encouragement, letter writing and the like during my education in philosophy. Not everything that they helped me attempt succeeded, but the help and backing have nonetheless been essential. They are (in order of appearance) Mikkel Gerken, Xavier Landes, Aaron James and Sune Holm.

During my three years as a PhD student, I have been lucky to have a great set of colleagues who it has been a pleasure to work with and to not-work with. I want to thank Klemens Kappel in particular for his efforts to create and sustain a great social environment that

spans professors, post-docs and PhD students. I am aware it is in no way a given that such an environment exists, and I believe it is extremely valuable.

I am also extremely grateful to have made some great friends during my time as a PhD student: My fellow PhD students Nana Kongsholm, Rikke Moresco Lange, Bjørn Hallsson, Tine Hindkjær Madsen and Josefine Pallavicini, and MA hang-arounds Katla Heðinsdóttir and Nanna Taszarek Holm. At lunches, Friday beer outings, nights out, Rome trips and dinner parties, I have spent more time with these people than most others during the last three years, and I have enjoyed it enormously. I have no plans of letting that end here.

I want to thank my office-mates, Nana and Tine, especially. Nana has been there from the very start, and I am sad that we could not (due to some unhappy circumstances) hand in our dissertations at the same time. I am extremely grateful to have had Nana to discuss the ins and outs of PhD life (and various more or less sensible philosophical ideas) with every step of the way – not to mention as a great co-teacher in our MA course on biotech ethics. Tine is probably the person I have discussed philosophy with most in the last three years. Although spontaneous hour-long discussions (not only about philosophy) in the middle of the day probably did not maximize my productivity, they were always enjoyable. I am especially grateful to Tine for persisting in the delusion that moral realism is true; for a while, showing Tine the error of her metaethical ways probably counted as one of my main hobbies. Although she won't admit it, I think I have managed to covert her to the true path.

Finally, and most importantly, I want to thank Ditte for always telling me I'm awesome whenever I have doubts, and for supporting me in the crazy project of becoming an academic. I couldn't have finished this thesis without her.

## Article overview

### **Article 1: Synthetic Biology and the Moral Significance of Artificial Life**

I discuss the moral significance of artificial life within synthetic biology via a discussion of Douglas, Powell and Savulescu's paper 'Is the creation of artificial life morally significant?' I argue that Douglas, Powell and Savulescu's definitions of 'artificial life' and of 'moral significance' are too narrow. Their definition of artificial life does not capture all core projects of synthetic biology or the ethical concerns that have been voiced, and their definition of moral significance fails to take into account the possibility that creating artificial life is conditionally acceptable. Finally, I show how several important objections to synthetic biology are plausibly understood as arguing that creating artificial life in a wide sense is only conditionally acceptable.

[Appears in *Bioethics*, 30(5), 372-379]

### **Article 2: Similarity Arguments in the Genetic Modification Debate**

In the ethical debate on genetic modification (GM), it is common to encounter the claim that some objection to GM would also apply an established, ethically accepted technology, and that this is a problem for the objection. I discuss how and to what extent this argumentative strategy, which I call a similarity argument, is useful. I construct a generic form of the similarity argument and show that it leaves some room open for the GM critic to avoid the seeming inconsistency that the similarity argument identifies. I then show how the GM critic can avail herself of this argumentative space in two specific cases, concerning two versions of the 'unnaturalness objection' to GM. Finally, I discuss the usefulness of similarity arguments in general.

[Accepted for publication in *Ethical Theory and Moral Practice*]

### **Article 3: Rationality, Thresholds and the Precautionary Principle**

The paper defends the precautionary principle from the charge of being irrational. Three irrationality-based objections are identified. The core of these objections is that the precautionary principle's use of thresholds of value and likelihood is unjustified, and that expected utility maximization is a preferable normative principle for risky choice. Against this, the paper argues (i) that thresholds of value are part of plausible and influential normative theories, and (ii) that the precautionary principle does not diverge more from ideal expected utility maximization than non-ideal expected utility maximizing procedures, and may do better in real-world choices.

[Under revision for *Ethics, Policy and Environment*]

### **Article 4: On the Cognitive Argument for Cost-Benefit Analysis**

Cass Sunstein has argued that the presence of a number of cognitive biases in our thinking about risks provide the basis of a 'Cognitive Argument' for the use of cost-benefit analysis in risk regulation as a corrective to the biases. I argue that one aspect of cost-benefit analysis, namely a scientific accounting of likely effects of regulation, can be justified as a corrective to cognitive biases, but that this aspect is not unique to cost-benefit analysis. The aspects of costs-benefit analysis that go beyond this depend for their justification on the desirability of 'coherence' in valuations of risks (i.e. that all risks of premature death are given a similar value). I argue that although there are methodological, instrumental and moral reasons to think that some form of coherence should be an uncontroversial goal of policy, cost-benefit analysis does not realize *those forms* of coherence. The use of cost-benefit analysis remains dependent on controversial moral ideas, and it cannot be justified as an a correction of cognitive biases.



# Introduction

## 1. Introduction

Synthetic biologists are currently working to make it possible to routinely engineer organisms in order to endow them with qualities that their human designers wish them to have. Synthetic biology thus means harnessing the powers of biology to a much greater extent than we have previously been able to by designing and fabricating artificial life forms that do things that are useful to us. Synthetic biology offers plenty of attractive possibilities, including contributing to developing solutions to some of the major challenges facing humanity in areas such as health, food, energy and environmental sustainability. But it also raises ethical questions. As I see it, two overarching questions have defined the discussion of synthetic biology from an ethical and societal perspective:

(I) Should we be engaged in the design and fabrication of organisms at all?

(II) How should synthetic biology be developed and regulated from a societal point of view?

The four articles that make up the core of this dissertation address aspects of these broad questions. Articles 1 and 2 concern question (I). Article 1 directly takes up the question of whether the creation of artificial life, as that activity is realized in synthetic biology, is a morally significant act. Article 2 discusses a common argument that aims to show that objections to designing and fabricating organisms (in synthetic biology as well as ‘traditional’ genetic engineering) should be rejected. Together, articles 1 and 2 provide an argument that the answer to question (I) should not be taken to be a *simple* yes, as many philosophers seem

to believe. I argue that many so-called intrinsic objections – objections to the activity of synthetic biology *per se* – make valid points if they are not interpreted as all-out rejections of the technology. Furthermore, I suggest that the relationship between intrinsic objections and questions of how society uses synthetic biology – i.e. between questions (I) and (II) – is closer and more complex than often thought.

Articles 3 and 4 concern one of the central issues falling under the heading of question (II), namely how the risks and uncertainties that characterise synthetic biology should be regulated. They deal with two of the most prominent approaches to risk management in environmental and health policy, namely the precautionary principle and cost-benefit analysis. Article 3 defends a moderate version of the precautionary principle from the charge of irrationality. Article 4 criticises an intriguing argument for the use of cost-benefit analysis, namely that it is a necessary corrective to the cognitive biases that mar our thinking about risk and uncertainty. Articles 3 and 4 thus collectively provide a tentative defence of the use of the precautionary principle and related strategies from criticisms mounted by proponents of expected utility maximization and cost-benefit analysis.

In this introduction, my main aim is to elaborate upon what I take the contribution of the four articles to be, especially with respect to providing answers to questions (I) and (II). This especially entails situating them in the general literature on the ethics of synthetic biology, and in aspects of the literature on environmental ethics and on the social management of risk and uncertainty. In addition, I will try to draw out some of the connections between the articles more clearly, and elucidate some of the methodological, theoretical and strategic choices that underlie my selection of the precise focus of each of the articles.

## **2. Synthetic biology and its reception**

Synthetic biology is a very broad and diverse field of research. In a report prepared by three scientific committees under the European Commission, synthetic biology is defined as “the application of science, technology and engineering to facilitate and accelerate the manufacture and/or modification of genetic materials in living organisms” (European Commission, 2014). This very broad definition emphasizes the technological aspect of synthetic biology, and this aspect is also the focus of this dissertation. However, it is worth noting that at least some researchers within the synthetic biology community have more purely scientific aims in mind, in particular a better understanding of the basic nature of life and how life may have originated (Holm 2013). The breadth of the field of synthetic biology has led to some discussion about whether there is a unified field at all, and to several competing classifications of the diverse projects that are going on under the heading of synthetic biology (Acevedo-Rocha, 2016; Bedau, Parke, Tangen & Hantsche-Tangen, 2009; Benner & Sismour, 2005; O'Malley, Powell, Davies & Calvert, 2008). I do not offer or endorse any particular classification, but instead provide some examples of projects and applications that have drawn attention from the public and/or exemplify important ethical issues raised by synthetic biology.

### **2.1 Synthetic biology research**

The most broadly publicised event within synthetic biology was the creation of a “synthetic cell” by scientists working at the J. Craig Venter Institute (JCVI) in 2010<sup>1</sup> (Gibson et al., 2010). The synthetic cell consisted of a chemically synthesized genome – a copy of a bacteri-

---

<sup>1</sup> A study of synthetic biology coverage in the media in Sweden and Italy from 2009 to 2013, found that around 1/3 of all articles with a “medium or strong connection” to synthetic biology were published in the three days following the JCVI’s announcement of the creation of the synthetic cell (Ancilotti & Eriksson, 2016).

al genome with the names of its inventors, among other information, added in a DNA code language – that was transplanted into an ‘empty’ cell of another bacterial species. The goal of this line of research is to be able to design and synthesize whole genomes and to insert them into empty cell ‘chassis’. Another important partial goal within this project is to create a minimal cell, i.e. a cell which has no genes except those that are essential for staying alive and reproducing. In 2016, the JCVI announced that this goal had also been reached (Hutchison et al., 2016).

Another important strand of research aims at introducing concepts of rational design taken from non-biological engineering domains into biotechnology (Endy, 2005). This includes the creation of an inventory of standardized biological parts, sometimes known as BioBricks, and efforts to make the design and creation of new biological systems less complex and divisible into smaller tasks. The aim is to make the design and production of biological systems easier and more efficient. This includes both construction of entire novel genomes – e.g. combining a minimal genome with all and only those ‘extra’ functions the designer wants – and the re-engineering of existing organisms, such as bacteria, yeast and algae.

The basic long-term aim of synthetic biology, including both of the strands of research just described, is to construct biological systems that have useful functions. Possibly the best-developed type of application is metabolic engineering (Keasling, 2010; Møller, 2014). Metabolic engineers modify organisms’ metabolic pathways in order to make them produce chemicals that have value for humans. The best-known example is the production of artemisic acid, a precursor to the antimalarial drug artemisinin, in yeast. Other examples (current and envisaged) include flavourings and fragrances (e.g. vanillin, the main flavour compound in vanilla, and the sweetener steviol), biofuels, rubber, and oils (e.g. as substitutes for environmentally-damaging palm oil in soaps). Using metabolically engineered

organisms, chemicals can be produced using only sugar or (in the case of algae and plants) sunlight as input.

Outside of chemicals production, engineered microorganisms are being developed to work in environmental protection and medical treatment (Khalil & Collins, 2010). With respect to environmental protection, organisms can be engineered to function as biosensors, i.e. to detect the presence of pollutants and toxic chemicals in the environment, such as arsenic in drinking water (French et al., 2011). Furthermore, they can be designed to 'clean up' such pollutants and toxic chemicals by metabolizing them into harmless compounds. Ideally, these two functions can be combined, yielding a biological search-and-destroy mechanism for pollutants. The use of engineered organisms in medical treatments follows a similar pattern: They can be designed to be used as diagnostics tools, detecting pathogens and other diseases, as treatments, providing targeted delivery of drugs, or as both at once (Jermy, 2011; Slomovic, Pardee & Collins, 2015).

The final application that I will mention is the use of synthetic biology as a tool in medical research and pharmaceuticals development. An example of the latter is vaccinations: the time required for developing a vaccine in response to an influenza pandemic can be cut from 6 months to less than a week (Dormitzer et al., 2013; Perkel, 2015). With respect to the former, synthetic biology can be used to better understand potentially dangerous pathogens, e.g. by synthesizing otherwise extinct pathogens, such as polio (Cello, Paul & Wimmer, 2002) and the H1N1 'Spanish' influenza (Tumpey et al., 2005). It can also be used to generate alterations in pathogens that would make them more dangerous, e.g. by making H5N1 influenza ('Bird flu') transmissible between humans (Osterholm & Kelley, 2012). This is known as 'gain-of-function research', and has the aim of enabling better prediction and monitoring of potential developing threat, and faster development of vaccines in case of a pandemic (Wimmer, Mueller, Tumpey & Taubenberger, 2009).

## 2.2 Reactions

As noted, the JCVI's announcement of the creation of an artificial cell generated much press coverage. It also prompted the president of the United States Barack Obama to order a report on synthetic biology from the Presidential Commission for the Study of Bioethical Issues (PCSBI, 2010). But engagement with synthetic biology from ethicists and policy makers began already in 1999, with a study funded by JCVI (Cho, Magnus, Caplan, McGee & Ethics of Genomics Group, 1999). Since then, a number of reports have been published by government agencies and bioethics think tanks (e.g. European Commission 2014; 2015a; 2015b; European Group on Ethics in Science and New Technology, 2010; Nuffield Council on Bioethics, 2012; Parens, Johnston & Moses, 2009). The ethical aspects of the technology have also been debated among synthetic biologists themselves through workshops, e.g. at the various instalments of the SBx.0 Conference Series,<sup>2</sup> and through the production of governance proposals (Church, 2004; Garfinkel, Endy, Epstein & Friedman, 2007).

The aspect of synthetic biology that has generated the most extensive reactions is research on pathogens. The synthesis of poliovirus was subject to a heated debate, with Craig Venter calling it “irresponsible” and “without scientific justification” (Pollack, 2002). Similarly, worries were voiced over the resurrection of the Spanish flu. Apart from the risk of accidental release – described by one expert as “almost a certainty” – the main discussion focussed on whether the genomic sequences potentially extremely dangerous pathogens should be published in journals or databases available to everyone (Von Bubnoff, 2005). The US National Science Advisory Board for Biosecurity (NSABB) decided to allow (*ex post facto*) the publication of the Spanish flu genome in 2005, but when researchers created a version of the H5N1 Bird flu that is transmissible between ferrets (and therefore likely between hu-

---

<sup>2</sup> <https://biobricks.org/programs/sbx-0-conference-series/>

mans as well) in 2011, NSABB intervened and the findings were published only in revised versions omitting crucial details (NSABB, 2012). An important reason for this difference in policy is that H5N1 is potentially much more dangerous than even the Spanish flu. The latter killed at least 20 million and perhaps as many as 100 million people, but had a mortality rate of at most 10% (Johnson & Mueller, 2002), while the former so far shows a mortality rate of up to 60%. Furthermore, the Spanish flu would likely be less dangerous than in the 1918 outbreak if it were to escape today.<sup>3</sup> A voluntary temporary moratorium followed, before the US government decided (partly in response to a near-escape of H5N1 from the Centers for Disease Control in March 2014 (Butler, 2014)) to halt funding for all pathogen gain-of-function research until safety procedures had been worked out (Reardon, 2014). The NSABB delivered its recommendations for such procedures in May 2016, with an aim to implementing them fairly soon (Kaiser, 2016).

With respect to the public, the main reactions to synthetic biology have come from a number of civil society groups that campaign for environmental protection and social justice. A consortium of these groups – led by the Friends of the Earth (FOE), the International Center for Technology Assessment (ICTA) and the Action Group on Erosion, Technology and Concentration (ETC Group) – has published a set of “principles for the oversight of synthetic biology” (FOE, ICTA & ETC Group, 2012). The main focus has been the use of chemicals derived from metabolically engineered microorganisms. Pressure campaigns have been organized against the use of synthetic biology based versions of vanillin<sup>4</sup>, steviol (Ribeiro & Thomas, 2015), artemisinin (Thomas, 2013) and algal oil in soaps (ETC Group, 2014). According to FOE, the vanillin campaign has resulted in several companies pledging

---

<sup>3</sup> A succinct statement of the NSABB’s rationale was given by a member of the board in response to a critical blog post: <http://www.virology.ws/2011/12/20/a-bad-day-for-science/>

<sup>4</sup> <http://www.foe.org/projects/food-and-technology/synthetic-biology/no-synbio-vanilla>

not to use synthetic biology vanillin.<sup>5</sup> In fact, the relevant companies have said that they would not use *any* artificial vanillin (the bulk of which is not produced using synthetic biology) – and in some cases this has been their policy all along (Watson, 2014). Despite increased coverage and the campaigns just described, the general public remains largely ignorant about the existence of synthetic biology, with 75% having heard just a little or nothing at all about synthetic biology according to a 2013 study (Hart Research Associates, 2013).

The worries and objections that philosophers, civil society groups and members of the public (when informed about its existence) voice concerning synthetic biology fall into three main groups: (1) Objections concerning the human relationship to life and nature; (2) Objections based on risk and uncertainty; and (3) Objections based on negative socioeconomic effects. I discuss categories (1) and (2) in the Articles. Discussion of those two categories is therefore assigned to the introductions to the Articles in the sections below. I do not, however, deal in any detail with objections falling into category (3) in the Articles. I therefore want now to discuss these objections briefly, and to explain how some of the content of the Articles bear on the socioeconomic objections.

## **2.3 Socioeconomic issues**

The main exponents of socioeconomic objections are the civil society groups. They repeatedly stress the potential negative effects that a move to biotechnology-based production of chemicals may have for those who currently produce these chemicals. In the case of vanillin, steviol and artemisinin, some of these producers are relatively poor farmers in developing countries, to whom loss of demand may be very harmful (Pollack, 2013; Ribeiro & Thomas, 2015; Thomas, 2013). Arguably these cases illustrate a more general problem of technological change leading to capital intensification and lower demand for labour, poten-

---

<sup>5</sup> <http://www.foe.org/news/news-releases/2014-08-haagen-dazs-says-no-to-synbio>



tially harming unskilled workers (Stiglitz & Greenwald, 2014). Furthermore, the civil society groups raise concerns about the environmental sustainability of a possible large-scale move to sugar-driven production of chemicals. Although this issue is high on anti-biotechnologists' agenda, it has not been discussed much in the philosophical literature or in ethical reflections of researchers. Insofar as socioeconomic issues are discussed, the main focus is on unequal access to new technology and on what intellectual property regime is desirable (Hunter, 2013; PCSBI, 2010, pp. 113-122 & 161-166; Schmidt et al., 2009, p. 5).

Paul B. Thompson (2012; 2015) argues that the lack of engagement with socioeconomic issues of the kind raised by the civil society groups is a serious failing in synthetic biology ethics (and bioethics more generally). He argues that synthetic biology should primarily be seen as a "platform technology" that enables a large number of different applications and commercial products. The field's primary effect on the world will thus likely be felt in the domain of the economy. Thompson suggests that this has already led to large speculative acquisitions of land in Africa by companies hoping to use it for biofuel production (Thompson, 2012, pp. 10-11). He hypothesises that neglect of socioeconomic issues is due to the fact that they are less 'sexy' than questions of safety, bioterrorism and boundary transgressions, and to the fact that the ethics of emerging technology is dominated by people and concerns that have their roots in *biomedical* ethics, where socioeconomic issues (excepting the issue of equal access) are not central.

I largely agree with Thompson, and my own neglect of the issues thus do not reflect a belief that they are unimportant. And some of the things I discuss in the Articles do pertain to socioeconomic issues. First of all, the notion of conditional acceptability that I discuss in Articles 1 and 2 can be applied to the socioeconomic conditions in which synthetic biology is embedded as well as the types of conditions I discuss (I will explain how in the next section when I discuss conditional acceptability). Second, Article 4 touches upon the distribu-

tion of benefits and risks, and engages with a framework for risk regulation that is essentially derived from economics. Third, risk plays an important role in normative theorising about economic issues, since the risk that individuals will be harmed economically is part and parcel of a dynamic economy. The significant role of public insurance in normative arguments for the welfare state reflects this fact (see e.g. Heath, 2011; Landes & Holtug, 2015). To the extent that my discussion concerns risk as a general issue (rather than only an issue in synthetic biology), it will have implications for socioeconomic issues as well.

### **3. Conditional acceptability, uniqueness and intrinsic objections**

My first two Articles engage with overarching question (I) as identified above, i.e. whether we should be engaged in the design and fabrication of organisms at all. My Articles reflect dissatisfaction with the way in which this debate has been conducted. I take issue with two common preconceptions about what an ethical evaluation of synthetic biology (or any technology) amounts to – namely that objections to and worries about synthetic biology in general (1) must provide reasons to reject synthetic biology *per se* as morally wrong or to impose general restrictions on the field, and (2) must pertain to features of synthetic biology that are *unique to* or *distinctive of* the field.

In this section I will introduce my first two Articles and explain how they challenge these two preconceptions. I will begin with (1) and move on to (2). Since objections falling into the first of the categories I identified above – that is, objections relating to our relationship to life and nature – are the main focus both for those who discuss question (I) in the literature and for my two Articles, I will review and comment on those objections as well.

### 3.1 Conditional acceptability

In the introduction to a collection of essays on synthetic biology and morality, Gregory E. Kaebnick and Thomas H. Murray suggest that an ethical evaluation of synthetic biology as such – as distinct from its possible consequences for human welfare – has three possible answers: That it is “morally troubling”, “attractive” or “morally neutral” (Kaebnick & Murray, 2013, p. 6). Elsewhere, Kaebnick understands critical views as claiming that synthetic biology is “intrinsically morally undesirable” or that it commits “intrinsic wrongs”, and he suggests that the ultimate upshot of critical views is that synthetic biology merits “special regulatory constraints” or even that it “should be banned” (Kaebnick, 2009). Bernhard Bartschi (2012a) interprets similar views as a “condemnation” of synthetic biology and synthetic biologists. My main claim here is that there is considerable room between the view that synthetic biology is morally troubling and the view that it is intrinsically wrong or that it should be banned or condemned.

My main contribution here is to suggest that the category of *conditional acceptability* is a useful possible alternative to the dichotomy between moral neutrality and moral wrongness with respect to general attitudes to synthetic biology. In Article 1, I discuss an argument made by Tom Douglas, Russell Powell and Julian Savulescu (2013) to the effect that creating artificial life is not “morally significant”. I understand this argument as relating to exactly the kind of general ethical evaluation of the activity of creating life that is taken up by my question (I). Douglas, Powell and Savulescu suggest that creating life is morally significant only if “there are moral reasons not to create artificial organisms, or factors that weaken our moral reasons *to* create them” (Douglas et al., 2013, p. 689). I suggest that, on this interpretation, creating life would be morally significant only if either (i) there were reasons to think that the fact that it constitutes creating life is a wrong-making feature of an action, or (ii) there were reasons to think that the whole set of actions falling under the heading of ‘creat-

ing life' are wrong. I then argue that conditional acceptability is a plausible alternative to (ii), i.e. that creating artificial life would be morally significant if there were reasons to think that the activity of creating artificial life is conditionally acceptable.

In Article 2, I argue that arguments based on the similarity between synthetic biology (or rather, genetic engineering more generally) and earlier technologies that are viewed as morally acceptable depend (among other things) on interpreting objections to genetic engineering as aiming at the conclusion that the technology is *not* morally acceptable. Otherwise there would be no rational pressure on the critic to bring her judgment concerning genetic engineering into line with her judgments concerning related technologies, such as selective breeding, since there is no *prima facie* conflict. But there are good reasons to think that critics, academic as well as non-academic, deem genetic engineering to be conditionally acceptable rather than simply unacceptable (or at least that the internal logic of their arguments support conditional acceptability at least as well as unacceptability).

By 'conditional acceptability' I have two closely related things in mind: (a) The negation of *unconditional* acceptability, and (b) the willingness to allow something only under certain conditions. In sense (a), conditional acceptability denotes the view that creating artificial life is not a type of action that falls within the domain of freedom. Read in this way, objections to creating artificial life provide a justification for limiting the liberty of would-be creators. At the very least, removing the creation of artificial life from the domain of freedom means placing those who want create artificial life under a standing obligation to provide justification for doing so (Gaus, 2005, p. 274). I believe this attitude well describes some public reactions to creating life, especially aversion to the *frivolous* creation of artificial life. For example, a focus group participant explained his view that we do not "have the right to genetically modify grains, animals" by asking: "if people don't like a certain color of cats, are they going to do that too?" (Hart Research Associates, 2014, p. 13). Similarly, Arthur

Caplan has suggested that people worry about “carefree, lighthearted, even irresponsible” creation rather than creation *per se* – i.e. that they object to *playing* God rather than to playing *God* (Caplan, 2009; 2010, p. 6).

In sense (b), conditional acceptability means accepting synthetic biology only if it is circumscribed in some way – i.e. accepting it only if certain conditions are in place. A standing demand for justification could count as a condition of this kind. In my Articles, I suggest a number of other conditions (some concerning genetic engineering more broadly):

- Ethics education for researchers;
- Continued ethical assessment of future developments, including public engagement;
- Adequate consideration of the interests of designed organisms;
- Supplementing technological solutions to societal problems with political and economic solutions;
- Effective risk management procedures, including precautionary approaches to some applications;
- Socioeconomic background conditions;
- Procedures that facilitate the freedom to choose whether one wants to consume genetically engineered products, such as labelling;
- That synthetic organisms are not used as replacements for natural ones.

The conditional acceptability view, and several of the conditions listed here, has been identified in qualitative studies of public opinion (Steurer, 2016). It can also be found in some

ethicists. Hans-Jürgen Link (2013) argues that worries based on the human relationship to life and nature would, if used to justify strong restrictions on synthetic biology, amount to invalid slippery-slope arguments, but that they do warrant continued assessment of future developments and a generally cautious approach. Anna Deplazes-Zemp (2012) suggests that objections based on biocentric views lead to the conclusion that creating life “should be performed under certain conditions or with a certain attitude”.

It is important to stress the difference between the view that synthetic biology is conditionally acceptable and the related view that the technology *per se* is morally neutral, while individual *applications* may be morally problematic (or morally obligatory, for that matter). First of all, not all of the conditional acceptability views listed have a close relative that says that some applications are acceptable and others not. For example, the view that genetic engineering products should be labelled is a view of this sort; as is the view about continued ethical assessment and (arguably) views about socioeconomic background conditions. Second of all, the conditional acceptability view is a view about what to think and do about synthetic biology (or another technology) *given* that some applications are good, some neutral and some bad from a moral point of view. While it is enormously valuable to preform ethical evaluation of each possible application, this gives no direct answers (except in the limit cases where all applications are bad or all applications are good) to the status of the technology *per se* – although there is a risk that an illicit jump to moral neutrality is made (e.g. Smith, 2013). But sometimes we do need to decide on an attitude to a technology *per se*. For example, we need to decide whether the technology should be funded by the public purse. In such cases, we need to be able to decide on how many of the bad aspects of a technology we will accept in order to get the good ones, and on how rigorous attempts to avoid possible bad aspects should be.

I do not mean to *defend* the application of all of the listed conditions to synthetic biology. Defending (or rejecting) even one of these would take much work, since each touches upon complicated issues such as the personal responsibility of scientists, the proper role of the public in decision-making concerning novel technologies, consumer sovereignty and socioeconomic justice. What I do argue is that each of these conditions merit such work, and that common objections to synthetic biology are more plausibly seen as arguments for the imposition of such conditions. Articles 3 and 4 can be seen as the beginnings of an adequate engagement with *one* of the conditions, namely what risk management regime should be implemented.

A final quibble about conditional acceptability is that it may relatively easily be misused, and arguably is misused by the civil society groups mentioned above (FOE et al., 2012). These groups have been accused of two types of misuse. First, that they dress up an “obstructionist” view to synthetic biology as a more moderate conditional acceptability view by letting acceptability depend on the fulfilment of conditions that are practically impossible to fulfil (Kaebrick, 2012). One type of impossible condition is an extremely strong version of the precautionary principle, which I will return to below. Another is large-scale change to the basic socioeconomic organization of the world; as Nathanael Johnson puts it, the groups “say people should be allowed to work with synthetic biology ... but only after we achieve utopia” (Johnson, 2014). Second, the groups have been accused of using the issue of synthetic biology as a mere foil for their real concern, which is to “challenge present societies by making generalized demands that are in direct contradiction with existing economic and social structures” (Lewontin, 2014). The issue of synthetic biology is thus used as leverage in a political campaign, and the relevant conditions are not particularly closely related to synthetic biology’s acceptability.

### 3.2 Uniqueness

In a much-discussed short article Joachim Boldt and Oliver Müller (2008) argue that “certain ethical implications of synthetic biology research go beyond those of genetic engineering”; in particular, they argued that synthetic biology marks the transition from humans manipulating organisms to our creating them. Christopher Preston (2008) argues that the organisms created by synthetic biology are the first properly unnatural or artificial ones, since they are not part of the great chain of evolution. On the other side of the debate, Erik Parens, Josephine Johnston and Jacob Moses (2008) respond to Boldt & Müller, arguing that the questions that synthetic biology raise – including all three of the categories mentioned above – are familiar from debates over other emerging technologies. Likewise, David Heyd (2012) argues that there is “nothing unique from an ethical point of view” in synthetic biology. And Beth Preston (2013) counters her namesake’s view by arguing that living artefacts have existed since the dawn of agriculture 10,000 years ago.

I have no quarrel with this discussion as such, but I suggest that a preoccupation with the question of novelty has led to an excessive focus on the novel aspects of synthetic biology – i.e. to the second preconception. The core of that preconception is the idea that objections to synthetic biology *per se* must pertain to features that are unique to or definitive of synthetic biology. This *uniqueness thesis* has the unfortunate consequence that it leads to a understanding of the essential feature of synthetic biology that is closely related to the ways in which it pushes the technological and scientific boundaries. In effect, this means understanding synthetic biology as essentially concerning the creation of organisms *de novo* (as opposed to manipulating existing organisms). Certainly, creation *de novo* is part of the project of synthetic biology, and could be described as the ultimate goal of the minimal genome and BioBricks projects. At the limit, the success of these projects would mean that scientists would have the ability to create organisms with all and only those functions they want (plus



the minimal genome necessary for the basic processes of life). But this is not the only, or even the main, aspect of the technology. We might just as well focus on the way in which synthetic biology brings engineering concepts into the biological domain, i.e. how it enables manipulation and creation of organisms that is easier, faster, cheaper, more efficient and more routinized. Such an understanding of synthetic biology would be much more conducive to asking questions like the socioeconomic ones Thompson wants us to ask, and less conducive to the more 'sexy' questions that concern humanity's relationship to life and nature. The engineering aspects of synthetic biology and the ethical issues that flow from them should not be ignored simply because they are not sufficiently unique.

In Articles 1 and 2, however, I largely follow the focus on synthetic biology as the creation of artificial organisms. But I challenge a narrower variant of the uniqueness thesis, namely that objections that pertain to life and nature need to target features that are distinctive of synthetic biology, especially relative to other biotechnologies. In Article 1, I criticise Douglas, Powell and Savulescu's definition of what constitutes creating artificial life. Much of my argument consists in conceptual analysis of 'artificial life'. I argue that Douglas, Powell and Savulescu's definition misses several plausible senses of 'artificial life', and that these senses are very much relevant in synthetic biology. From the point of view of the uniqueness issue, however, my main claim is that those aspects of the creation of artificial life that critics most plausibly object to are aspects that are also present in (at least) traditional genetic engineering, but that are merely exacerbated in synthetic biology.

In Article 2, I target a widespread argument that is based on, and further motivates, the uniqueness thesis. According to this argument, objections to genetic engineering based on facts such as that genetic engineering is unnatural or that it amounts to the production of living artefacts are threatened by the fact that earlier and universally accepted technologies also possess those features – for example, as it is exceedingly often observed, that domesti-

cation and cultivation of animals and plants also alters those species at the genomic level. I broaden the scope from synthetic biology to genetic engineering more generally because that is the context in which the argument is most common. Furthermore, many of those who are critical of synthetic biology are also critical of genetic engineering more broadly, and for the same reasons. My main argument is that there is room for maintaining that genetic engineering is problematic though (for example) domestication is not, even if the feature to which one objects is also present in domestication; and that subtle differences in the meaning of (for example) naturalness matter.

### **3.3 The human relationship to life and nature**

The final contribution of my first two Articles is to the discussion of objections to synthetic biology (and genetic engineering more generally) that are based on the human relationship to living beings and the natural world. In my Articles, I mainly defend such objections from certain attempts to dismiss them. My main point is that the relevant arguments are too superficial in their treatment of these objections, and that they must be countered individually and with attention to the details of the points made. I will return to how the arguments discussed in the Articles misinterpret the objections at the end of this section. But first I want to practice what I preach and engage these objections in a little more detail. I will suggest that none of them provide reasons not to create artificial organisms.

We can distinguish three rough categories of objections of this kind: Those concerning the *moral standing* of artificial living beings; those concerning the *intrinsic value* of artificial living beings; and those concerning the *attitudes* to life and nature that creating artificial living beings implies.

### 3.3.1 Moral standing

A being has moral standing just in case its interests matter morally and it can (therefore) be wronged. The creation of artificial living beings raises two questions with respect to moral standing: (i) Does artificiality matter for whether a living being has moral standing? (ii) Does the creation of artificial living beings constitute wronging them? With respect to the first question, the majority view is that artificiality does not matter for the moral standing of living beings (Attfield, 2012; Baertschi, 2012b; Deplazes-Zemp, 2012; Huesken, 2014; Sandler, 2012). The general reason is that moral standing is thought to be grounded in intrinsic properties of the being, such as sentience or its 'having a good of its own', which are not affected by the artificiality of the organism.

One exception to this general consensus is Christopher Preston (2013a, p. 122-123).<sup>6</sup> Preston gives two reasons to think that artificial organisms have no, or only a diminished, moral standing. First, he argues that the fact that the products of synthetic biology are *organisms* is only a "concomitant" or "incidental" attribute of them. The purpose for which the organism was designed, on the other hand, is a necessary feature of the entity. For example, an organism designed to capture and bind carbon in the atmosphere is, on Preston's view, necessarily a carbon-capturing device, and only incidentally an organism. Therefore, "the sense in which synthetic bacteria are tools serving a particular purpose takes precedence over the sense in which they are autonomous organisms" (Preston, 2013a, p.123). An obvious objection is that in a case where a *person* performs some function, such as checking tickets in the subway, we may not treat him as a *mere* tool. Preston counters that the difference

---

<sup>6</sup> Preston argues that artificial organisms have less "intrinsic value" than natural ones, but he does not clearly distinguish between moral standing and intrinsic value in my sense. The reasons described here seem to me to pertain mainly to moral standing, while the others mainly concern intrinsic value. I return to the latter below.

can be explained by the fact that being a ticket checker is incidental to the *very existence* of the person, while it is not incidental to the very existence carbon-capturing bacterium.

Second, Preston argues that artificial organisms have a “bifurcated teleology”. On the one hand, it has an internal, “organismal” teleology like natural organisms do; on the other hand it has an external, “artefactual” teleology imposed by the designer and defined by her goals and purposes in designing the organism. Moral standing is only conferred on the basis of organismal teleology – entities that possess only artefactual teleology, such as thermostats, are typically not thought to have moral standing. Since artefactual teleology is primary in artificial organisms, their moral standing is at least diminished relative to natural organisms. To bolster this claim, Preston suggests that the fact that people are willing to treat domesticated animals much worse than wild animals suggests that they take domesticated animals (i.e. partially artificial organisms) to have lower moral standing than wild animals (i.e. purely natural organisms).

Neither of these arguments seems to me plausible. To take the second one first, the argument relies on seeing the two kinds of teleology as being in conflict, so that an entity cannot be simultaneously fully organismically-teleological and fully artefactually-teleological. I see no reason to accept this view. Furthermore, the relevance of the maltreatment of domestic animals is questionable. For one thing, the view that the welfare of such animals is inconsequential seems to me to be very much a minority position today. For another, even if people do think domestic animals’ welfare does not matter, they may simply be wrong. Furthermore, it is doubtful if those who *do* deny the importance of domesticated animals’ welfare also think that *wild* animals’ welfare is important.<sup>7</sup>

---

<sup>7</sup> It is true that some environmental ethicists, in particular J. Baird Callicott (1980), argue that domesticated animals have less value than wild animals. But even Callicott does not defend a differential moral standing for wild and domestic animals, since he (i) initially (in the 1980 paper) denies moral

Concerning the first of Preston's arguments, the explanation of the difference between the carbon-capturing bacterium and the ticket-checking person does not seem to me adequate. Even in cases where a person is (or would be) *created* for a purpose – and hence that purpose would be essential in explain why the person exists – we would not be justified in treating that person as a mere tool. The worry that such persons may *in fact* end up being treated as mere tools underlies well-known objections to 'saviour siblings' – children born to provide biological materials that could save an ill older brother or sister (see Boyle & Savulescu, 2001). If the intuition that saviour siblings may not be treated as mere tools were not very solid and very widespread, this objection would not have the considerable bite that it does. This suggests that the general consensus position described above, according to which intrinsic features of entities ground moral standing, is on the right track.

Even if artificiality does not affect moral standing, the possibility of artificial life does present a puzzle for some biocentrist views. If moral standing is grounded in the fact that an entity is teleologically organized, as some argue, then either it follows that traditional artefacts (such as Preston's thermostat) have moral standing, or a distinction must be made between the kind of teleology possessed by *living* beings and that possessed by artefacts. The standard way of trying to do the latter is to appeal to a specifically *biological* teleology that is grounded in the process of evolutionary selection. But at least potentially, some artificial organisms will lack an evolutionary history, and hence biocentrists lack the resources for arguing that living beings, but not balls and chairs, have moral standing (Holm, 2012b). Some biocentrists have bitten the bullet and accepted the implication that ordinary artefacts also have moral standing (Basl & Sandler, 2013).

---

considerability to *individual* wild animals as well, and (ii) later softens his position considerably on both domestic and wild animals (see Lo, 2010).

The second issue concerning moral standing is whether the act of creating artificial life constitutes wronging the organism. From the point of view of 'mainstream' biocentrism, the answer seems to be that it depends on whether the quality of life of the created organisms is unacceptably low. Perhaps the 'onco-mouse', engineered to develop cancer, would count as a being that is wronged in this way (Attfield, 2012, p. 86). Even in cases such as this, where the way an organism is designed results in predictable suffering, some might deny that the organism is being wronged. A version of the non-identity problem arises, since the individual onco-mice could not have existed with a higher quality of life. I will refrain from entering this complicated debate. For current applications of synthetic biology, a more immediate problem is that microorganisms' interests are so rudimentary – perhaps restricted to self-maintenance only – that it is difficult to see how any designed characteristic could conflict with them (Link, 2013). Furthermore, current applications only pertain to microorganisms, and the view that we are wronging created organisms thus requires a form of biocentrism that only a small minority of philosophers (and probably even smaller minority of lay people) would accept.

Less mainstream biocentrists do suggest that the very act of creating artificial life constitutes wronging organisms. The main proponent of such a view in the context of synthetic biology is Joachim Boldt (2013a; 2013b). Boldt argues that taking an engineering approach to life, as synthetic biology does, necessarily entails viewing living beings in mechanistic cause-and-effect terms, rather than as purposeful beings striving to achieve their own good. Taking this view, in turn, precludes seeing created organisms as having moral standing. Designing and creating an organism thus constitutes denying that it has moral standing. And according to Boldt, this will also apply when created organisms are sentient animals or even human beings rather than mere bacteria.

I doubt that the basic premises of this argument are true. The mechanistic ('reductive') and the purposeful ('holistic') views of organisms are complementary rather than mutually exclusive, and it is doubtful that synthetic biology takes a thoroughly mechanistic view of life (Lewens, 2013; Powell, 2015). Even if they were mutually exclusive, taking the mechanistic view would only undermine moral standing if moral standing were tied to the purposefulness view. At least some bases for moral standing, such as sentience, are not dependent on that view. At the limit, Boldt's line of thinking seems to reduce to the question of the compatibility between freedom (a clear basis for aspects of moral standing) and deterministic explanation. While this is a vexed debate, it is safe to say that it is an open possibility that a mechanistic view even of human beings is compatible with seeing them in purposeful terms and grounding moral conceptions on purposefulness.

### **3.3.2 Intrinsic value**

The second category of objection concerns the *intrinsic value* of organisms. Intrinsic value, as I understand the term here, is distinct from moral standing in that it does not depend on whether the entity can be wronged. A great work of art can have intrinsic value – it can be good in itself that it exists – but it does not plausibly have moral standing. Several authors have argued that *natural* living beings possess a value that *artificial* living beings lack. Perhaps the most thoroughly worked out version of this view is Keekok Lee's (1999). On Lee's definition, an entity is an artefact if, and to the degree that, it "embod[ies] human ends" (Lee, 1999, p. 37). Artificiality is thus a degree term for Lee. The main determinant of the degree of artificiality is the 'deepness' of the technology used, which for Lee increases as a technology intervenes at more fundamental levels. The ultimate artefact would be one that was built directly from fundamental matter (i.e. atoms, subatomic particles, strings or whatever turns out to be at the ground-floor of physical reality). The implication for the domain

of biotechnology is that life forms that are the product of human interventions can be ranked according to their level of artificiality: Traditional trial-and-error selection produces organisms that are less artificial than those resulting from scientifically informed breeding and hybridization, and these are again less artificial than organisms created through genome-level interventions – including both traditional GMOs and the products of synthetic biology (Lee, 1999, p. 52). Lee’s claim is now that naturally occurring entities have “independent value” – a value that is grounded in the *very fact* that natural entities are independent of humans and human intentions<sup>8</sup> (Lee, 1999, p. 178). Consequently, organisms created by biotechnology, and especially organisms created *de novo*, lack a value that natural organisms possess.

Christopher Preston (2008; 2013a; 2013b) has adapted Lee’s argument to synthetic biology. Initially (in the 2008 paper), Preston rejects Lee’s explanation of the special value of natural organisms in terms of independence (largely by reference to the similarity between biotechnology and selective breeding, as discussed in Article 2). Instead, he suggests that the fact that artificial organisms are not connected to the historical process of evolution robs them of a value possessed by natural, evolved organisms. On this early view of Preston’s, the products of synthetic biology *uniquely* lack value, since other manipulated organisms (such as those who have been selectively bred or genetically modified in the traditional way) retain their connection to the evolutionary process. Later (especially in 2013a), Preston accepts that lack of independence in Lee’s sense can also at least diminish the intrinsic value of organisms (Preston, 2013a, pp. 119-120). He also seems to endorse the view that domesticated animals lack part of the intrinsic value that natural organisms have, although it is

---

<sup>8</sup> It is worth noting that this idea has a certain paradoxical flavour: If it is important that value is not anthropocentric, then it seems strange to suggest entities’ relationship to human being is of such great importance. From the point of view of the non-humans themselves, it presumably matters little whether they are independent (in Lee’s sense) or not.



unclear whether this lack pertains to intrinsic value, moral standing, or both (cf. note 6 above). At any rate, on Lee's and Preston's views, artificial organisms lack intrinsic value, or have less intrinsic value than natural organisms.

Their lack of intrinsic value may mean that we lack one kind of reason for creating new life forms and for preserving them if they were to become threatened with extinction – artificial life forms, in other words, do not contribute to valuable biodiversity (Boldt, 2013c). Could the lack of value also be a reason to *refrain* from creating artificial life? Lee argues that deep technologies are “nature-replacing”, and that they “transform the natural into the artefactual” (Lee, 1999, pp. 6-7). In the most straightforward reading of these claims, it is easy to see how a problem might arise. If synthetic biologists (and other biotechnologists) are literally replacing valuable natural entities with non-valuable artificial entities, then a loss of value would result. But this interpretation does not warrant an objection to creating artificial organisms *per se*, but only to substituting natural for artificial organisms. An argument of this type has been made against restoration of natural ecosystems, which has been denounced as “faking nature” (Elliot, 1982) and a “big lie” (Katz, 1992). But at least currently, synthetic biology does not lead to there being less natural organisms in the world, as Preston concedes in his later paper (2013a, pp. 112). This interpretation of the nature-replacing objection seems at best to concern what might happen in the far future (and Lee's book is strewn with expressions such as “could ultimately”, “in the long run” and “would eventually”). Taking this to be an objection to current synthetic biology is to commit a slippery slope fallacy.

### **3.3.3 Attitudes**

But Lee sometimes seems to suggest that the mere fact that it becomes *possible* to replace natural entities with artefacts is sufficient (Lee, 1999, pp. 117-120). On this interpretation, the

possession of a technology that allows us to substitute artefacts for natural entities puts humanity in a position of *dominance* over nature, in the sense that it is up to us to decide whether there should exist natural entities. This takes us to the final category of objection, those concerning the *attitudes* to life and nature that creating artificial living beings implies. The typical attitude that synthetic biology (and other technology) is claimed to embody is that of total mastery and subordination of the natural world to the desires of human beings. Boldt & Müller (2008, p. 388) suggest that creating organisms *de novo* amounts to seeing nature as “a blank space to be filled with whatever we wish”. A weakness for views of this type is that their connection to action is less than obvious. As Boldt (2013b, p. 40) concedes, this view is less about acts and more about “goodwill, attitudes and virtues”. Hence it would be wrong to assume that these views directly provide reasons for or against certain actions.

Presumably, however, the implication is that we should not *want to be* the kind of people who dominate nature (and that we should want to avoid creating organisms, at least *de novo*). The question then is *why* we should not want to be this kind of people? One possibility is that the attitude of dominance *just is* objectionable or undesirable. Lee argues that it amounts to narcissism and egomania, and therefore “either pathological or immature” (Lee, 1999, p. 201). A second possibility is to argue that such an attitude is inconsistent with granting living beings moral standing or intrinsic value. As I argued above, Boldt seems to hold that taking a certain attitude to living beings – the mechanistic view – precludes seeing them as having moral standing. Lee analogously suggests that the attitude of dominance and its attendant striving for control “necessarily involves trampling on the legitimate ethical demands of nonhuman others” (Lee, 1999, p. 202). A third possibility is that the attitude is detrimental to *our own* interests. Lee laments the prospect of a world where “[w]herever one turns, one only sees images of oneself; wherever one shouts, one hears only one’s own

echoes." In such a world, humanity becomes "lonesome" and "imprisoned within an existential or ontological solipsism of its own making" (Lee, 199, p. 194). G.A. Cohen (2011, p. 207) takes a similar view, arguing that "the attitude that goes with seeking to shape everything to *our* requirements ... contradicts our own spiritual requirements" and is "repugnant, and, at the limit, insane". A fourth possibility is that the attitude of mastery, even if only taken to nonsentient life, may lead to diminished respect for human or sentient life (Boldt, 2013b, p. 44). This is analogous to Kant's argument against animal cruelty, but moved one branch down the phylogenetic tree (Link, 2013, p. 441). A fifth and final possibility, similar to the fourth, is that the normative attitude of mastery may lead to an exaggerated *factual* belief in our ability to control and predict the result of interventions in the natural world. The attitude-based objection is thus linked to risk-based concerns (Boldt 2013a, pp. 400-401; 2013b, pp. 44-47).

As I argue in Article 1, I do not think investigating the internal mental states of synthetic biologists can dispel (or confirm) the attitude-based objections. The objections do not (or do not only) concern such mental states, but rather what Boldt, borrowing a phrase from Hans Jonas, calls "the basic precepts of one's doing" (Boldt, 2013b, p. 40). The question is whether the attitude of mastery is necessary to make sense of the action of creating artificial life. Even if we take this view of attitudes, however, it is doubtful whether creating artificial life in the way synthetic biologists currently do only makes sense if in the light of an attitude of mastery that is narcissistic, or that is inconsistent with granting moral status or intrinsic value to living beings. First of all, the attitude-based objections mainly target the project of designing and creating life *de novo*. But much of what is going on within synthetic biology does not amount to creating life *de novo*, and as such does not (on the objection's own terms) exhibit the attitude of mastery (Thompson, 2012, p. 5). Second, even those aspects of synthetic biology that best fit the paradigm that Boldt, Lee and others have in mind

do not require the attitude of mastery. Tim Lewens (2013) argues that the engineering approach to life that synthetic biologists undoubtedly take is better explained as a set of pragmatic ways of dealing with the complexity of organisms by simplification than as the reflection of a thoroughgoing rational-design view of life (i.e. Boldt and Müller's "blank space" view of nature). Similarly, making sense of a person that designs and creates novel types of microorganisms does not require ascribing to him a lack of recognition of any intrinsic value or moral standing in natural life. At most it requires ascribing to him the belief that manipulating non-sentient living beings does not violate any valid moral norms. Much of the attitude-based objection thus reduces to the question of whether it is morally objectionable *for other reasons* to do what synthetic biologists are doing.

A second problem for most of the explanations of why the attitude of mastery should be avoided is that they rely on some connection, causal or otherwise, between the attitude of mastery and something more obviously problematic. But the existence of this connection can in every case be questioned, at the very least when the attitude is appropriately limited to what is necessary to make sense of creating life as synthetic biologists are now doing. There is room between even an attitude of mastery, if this concerns only non-sentient life, and attitudes that are narcissistic and egomaniacal. Likewise this kind of mastery falls short of "seeking to shape everything to our requirement", not to speak of *achieving* anything like the situation where we see only our own mirror-image in the outside world. With respect to the two final possible reasons why the attitude of mastery is problematic, it seems to me unlikely that our creating artificial life at the microorganismic level will lead either to a lack of respect for higher life-forms or problematic approaches to risk assessment and risk management. Those who put forward these worries have done little to argue that the effects *are* likely (Link, 2013, p. 445). Boldt (2013b, p. 47) in effect suggests that his view depends on

seeing the desire to create microorganisms as *already constituting* a general lack of respect (and concedes that some will not share this analysis)

There is another, more pragmatic variant of the attitudes-based objection, namely that synthetic biology is motivated by the same ideology that has created the problems synthetic biology is supposed to solve. This idea is common in the literature on the ethics of agricultural biotechnology (i.e. the use of genetically modified crops in agriculture). Ronald Sandler (2003, 2007) has argued on this basis that there is a presumption against the use of GM crops. The general claim is that GM agriculture is an expression of an attitude to nature based on a desire to control and master it, rather than on humility and respect. Unlike Lee and Boldt, he attempts to define this attitude more precisely, as the disposition “not to adapt [our] lifestyles to the earth, but to adapt earth to [our] lifestyles” (Sandler, 2004, p. 310). On Sandler’s view, this is not a reason to be against GM agriculture *tout court*, but only a reason to object to relying of GM as a quick technological fix to the major challenges humanity faces. Instead, we should at least *also* look to solve those problems through political and behavioural reforms (for an overview of the ‘technological fix’ critique of biotechnology, see Scott, 2011).

This line of argument reveals a very common, but often overlooked, aspect of critical views on biotechnology, namely that they call for consideration of (and preference for) *alternatives* to biotechnological solutions (rather than rejection of biotechnology altogether). Taken in one way, this call for considerations of alternatives seems entirely reasonable; we should consider all possibilities when it comes to solving complex problems such as ensuring environmental sustainability or creating an adequate food system for a growing global population, and we should avoid framing challenges that are (also) social and economic as (solely) technical ones. But if the view is that technological solutions should only be considered when all else fails, then it must rely on independent arguments to show that we have

even a *pro tanto* reason to avoid using biotechnologies. It would thus be an only slightly more moderate version of an anti-technology view.

#### **4. Risk, uncertainty and precaution**

I now turn to question (II), i.e. how synthetic biology should be developed from a societal point of view. The notion of conditional acceptability suggests that the gap between this and question (I) is less marked than it might seem – the conditional acceptability view, in a very broad variant, says that the acceptability of doing synthetic biology depends on whether it is developed in certain ways, and on what arrangements are necessary to secure such a development. Articles 3 and 4 contribute to the discussion of what arrangements we should instate in order to secure a development that is satisfactory with respect to the potential adverse effects of synthetic biology on human health and the environment, i.e. with respect to risk and uncertainty.<sup>9</sup> My Articles mainly deal with generic principles for risk management – in particular the precautionary principle and cost-benefit analysis – rather than with the specifics of what risks and uncertainties synthetic biology creates. In this section, I therefore first want to give an overview of how risk issues have been dealt with in the synthetic biology literature, and to explain how this relates to my choice of focus in the two Articles. I will then introduce each Article in turn.

While much of the philosophical literature on synthetic biology stresses that risk is the main ethical issue, few actually say anything very specific about what the risk-related issues are and how they should be handled. The typical claim is merely that risk is not sufficient to make warrant a wholesale rejection of synthetic biology. The main exception to this

---

<sup>9</sup> Strictly speaking, risk and uncertainty are also as important with respect to adverse *socioeconomic* effects. However, the arrangements that may be needed to make socioeconomic risk and uncertainty acceptable is a very different from the ones that are relevant in the case of health and environmental risks. When I talk about risk I will therefore mean risk to health and the environment only.

trend is the dual-use issues that are a result of research on pathogens, such as the H5N1 and polio research discussed above. The dual-use issue arises when research that has considerable benefits may also be intentionally misused. By far the most discussed issue is the use of synthetic organisms for biological warfare or terrorist attacks, but more mundane dual-use problems exist as well. For example, opioids can now be produced in engineered yeast, which makes it possible to produce painkillers such as morphine in an easier and cheaper way (DeLoache et al., 2015; Gelanie, Thodey, Trenchard, Interrante & Smolke, 2015). But it also makes it possible to produce drugs such as heroin equally easily and cheaply, which has led to some debate on this application, including a self-imposed moratorium from one of the teams working on this (see Ehrenberg, 2015; Service, 2015; Yang, 2015).

Developments in synthetic biology – significantly aided by the post-9/11 anthrax letters incident – has led to an increased attention on the dual-use problem in general terms (Miller & Selgelid, 2007; National Research Council, 2004; Rappert & Selgelid, 2013). Synthetic biology creates (or exacerbates) three aspects of the dual-use problem.<sup>10</sup> First, it promises to make engineering of organisms easier and cheaper, thus making it easier and cheaper to produce pathogens and (as the morphine example shows) perhaps easier to produce possibly harmful substances via metabolic engineering. The rise of do-it-yourself biology (or garage biology, or biohacking) intensifies the worry that dual-use biotechnology becomes available to people who are less reliable and surveillable than scientists at major universities and companies. Second, increased proficiency in designing organisms raises the possibility of ‘designer pathogens’ that are intentionally made more dangerous or useful as weapons than similar natural pathogens. This prospect, especially, has landed synthetic biology on a list of “12 risks that threaten human civilization” prepared by researchers at

---

<sup>10</sup> The factual claim that synthetic biology does create/exacerbate dual-use problems have been challenged by Jefferson, Lentzos & Marris (2014).

the Future of Humanity Institute and Martin School at Oxford University (Pamlin & Armstrong, 2015, p. 109). Third, as discussed above, the publication of the genomes of known pathogens place the blueprints for a potentially devastating weapon in the public domain.

The third issue (often in combination with the first) has been the most prominent synthetic biology risk issue discussed in mainstream bioethics. Douglas and Savulescu (2010) calls for the development of an “ethics of knowledge” that can guide us in determining whether to publish information that significantly increase the ability of evil-doers to produce dangerous pathogens. It is at least a possibility that such an ethics of knowledge ends up demanding some form of more or less restrictive pre-censorship of sensitive scientific publications. Michael Selgelid (2007) argues for a similar position, in particular for stricter formal regulations and against self-regulation as a solution. The idea of restricting the freedom to publish scientific results has met with some criticism. Kevin Smith (2013, pp. 461-462) argues that restrictions are unjustifiable on consequentialist grounds, since “it is not possible *a priori* to discern those pieces of scientific knowledge that will generate disutility, cf. those that are either neutral or generative of beneficial consequences.” And Robin Pierce (2012) argues that any censorship scheme raises questions of who decides what research should be censored and of prioritizing the interests of those who want to feel safe over those who stand to benefit from the relevant research.

Outside dual-use, the risk-related problems generated by synthetic biology have received much less attention from philosophers. A minor exception is gain-of-function research more generally, which has been subjected to an ethical analysis in connection with the NSABB’s general analysis of safety that followed the H5N1 studies in 2014 (Selgelid, 2016); but that analysis proceeded from the assumption that “[gain-of-function research] is a subset of “dual-use research””. In fact it seems to me that the opposite is true – dual-use is *one* problem that gain-of-function research generates, namely that of intentional misuse of



‘enhanced’ pathogens. The other major problem is the unintentional release of such pathogens. Most of what has been written about dual-use, especially for and against censorship of science, has no bearing on the problem of unintended release. Arguably censorship, which is a possible solution to dual-use, is counterproductive with respect to unintended release, since the availability of information about the escaped pathogen can be used to combat it. Furthermore, dual-use raises some unique questions that do not apply to gain-of-function research generally – especially questions concerning the responsibility of scientists, funding bodies and others for wrongs that others perpetrate. I will not discuss these dual-use-specific questions, but instead focus on more generic risk management issues. Such issues are at least *also* of central importance in the context of dual use. Even though dual-use pertains only to cases where harm is due to some intentional (criminal) action, it shares the general structure of other choices under risk and uncertainty when analysed from the point of view of scientists, funding bodies and regulators – for example, Douglas (2013) argues that expected-value theory is apt for handling dual-use as well as more standard cases. Furthermore, the precautionary principle and cost-benefit analysis (and their relative merits) have been discussed in the context of dual-use (Clarke, 2013; Kuhlau, Höglund, Evers & Eriksson, 2011).

Even though philosophers have been more or less absent from the debate, the question of how synthetic biology should be regulated with respect to its attendant risks to health and environment has been discussed, especially by practitioners themselves and by lawyers. The main issue here is whether synthetic biology warrants or requires *specific* regulation, or whether it can instead be covered by existing regulations of similar products. Civil society groups have called for “synthetic biology-specific regulations”, and argued for a moratorium until such regulations are put in place (FOE et al., 2012, p. 4). On the other side

are those who argue that we should ‘regulate the product, not the process’, i.e. that synthetic biology products should be subject (only) to those regulations that cover similar products.

The process/product debate with respect to synthetic biology has a peculiar second-order nature to it, since it piggybacks on a similar debate relating to traditional genetic engineering. The European Union has instated process-regulation of GMOs, and synthetic biology is regulated by those regulations. In one sense, then, synthetic biology is regulated as similar products, but since those products are process-regulated, synthetic biology is ultimately process-regulated. The contrast case is regulation in the United States, which is thoroughly product-based. Products of genetic engineering, including synthetic biology, are regulated according to what type of products they are. Thus synthetic biology-based drugs are regulated by the Food and Drug Administration’s Federal Food, Drug and Cosmetics Act (FDCA); synthetic organisms used as pesticides fall under the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) under the Environmental Protection Agency; and industrial use (such as chemicals production using metabolic engineering) is covered by the Toxic Substances Control Act, also under EPA administration (for an overview of regulatory regimes see Bar-Yam et al., 2012; Bergeson et al., 2015; Rodemeyer, 2009).

I am generally supportive of regulating product rather than process. But regulating product rather than process does not in itself mean that no new regulations are necessary, since synthetic biology could produce products that are relevantly different from seemingly similar products of traditional genetic engineering and other technologies. Two aspects of synthetic biology are especially interesting, since they increase the level of uncertainty of these products relative to existing similar products. First, synthetic organisms in non-contained use, e.g. as pesticides or for bioremediation, may be able to survive, reproduce and evolve. This is a relevant difference from chemical and mechanical methods that are typically used for similar purposes, since their effects will be less predictable, and any nega-

tive impact may be more persistent. (However, organisms are already used in such capacities, e.g. the use of the *Bacillus thuringiensis* (of GM crop fame) as a pesticide, so there is some relevantly similar products in existence). Second, as synthetic organisms become more and more different from any natural counterpart, risk assessment becomes harder. The current risk assessment procedure of GM crops in the US relies on the concept of “substantial equivalence”, whereby a GM crop is considered safe if it is sufficiently similar to an existing crop. This concept could plausibly be extended to some synthetic biology products, such as yeasts that produce vanillin rather than alcohol. But for organisms that have no close natural counterpart, we cannot use this source of knowledge in risk assessment. *Prima facie* this strengthens a case for pre-release testing of synthetic organisms that do not correspond closely to existing organisms. Furthermore, the substantial equivalence doctrine is itself controversial, especially with regard to procedures for deciding when one organism is substantially equivalent to another (McGarity, 2002a).

The focus of Articles 3 and 4 reflect this mix of novelty and familiarity with respect to regulation of synthetic biology. Article 3 deals with issues that arise when there is significant uncertainty about the effects an application might have and how likely possible negative effects are. The Article defends the use of the precautionary principle in such cases, especially against challenges from proponents of expected utility maximization procedures. Article 4 reflects the fact that the risks of synthetic biology are (and in many cases should be) regulated in the same way as other activities that create risks to health or the environment. A common tool in such regulation is cost-benefit analysis, which combines scientific and economic analysis to arrive at an assessment of overall costs and benefits of different regulations from a societal point of view. I criticise one argument in favour of using cost-benefit analysis in this context. Common to both Articles is a focus on defences of expected utility maximization and cost-benefit analysis – which are closely related and underlies

what might be described as mainstream risk management – that attack alternatives for being *irrational*.

#### **4.1 The precautionary principle**

The debate around the precautionary principle is extensive and points in many different directions. It has aspects that fall within epistemology, environmental philosophy, political philosophy, ethics and the philosophy of science. In Article 3, I deal with the precautionary principle as a principle that is supposed to guide actions in policy-making and legal contexts. My point of departure is the account of the principle defended by Daniel Steel (2014). The basis of Steel's account is what he calls the "tripod". This interpretation of the principle is widely adopted, and it has been suggested that it can be found in more or less all formulations of the principle that exists in official documents (Randall, 2011; Sandin, 1999; Trouwborst, 2006). The tripod analysis takes the principle to be a general schema, according to which the existence of some level of *evidence* that an activity may lead to some *harm* triggers the demand for a precautionary *action*. Steel calls these three elements the *knowledge condition* (K), the *harm condition* (H) and the *recommended precaution* (P). K and H set *thresholds* for evidence and harm, respectively: If both thresholds are met, the demand for action is triggered. Fixing the variables yields a *version* of the precautionary principle. In fact, P is rarely specified to any detail in official statements of the principle – although it is subject to some generic conditions such as that it must be effective, proportionate and consistent with recognized legal principles such as non-discrimination (European Commission, 2000; Trouwborst, 2006, Ch. 5). This suggests that the relationship between H and K, on one side, and P, on the other, is primarily a *negative* one, i.e. that policies that trigger an H-K combination are *prohibited* (with other criteria needed to select among permissible policies). Such an analysis is suggested by Sandin (1999, p. 894), and by Martin Peterson's interpretation of the

principle as a “transformative decision rule” that removes certain options from consideration (Peterson, 2003). At least in his formal appendix, Steel adopts this approach as well.

The argument I put forward in Article 3 flows from engagement with Steel’s discussion of Cass Sunstein’s objection that any non-trivial precautionary principle is *incoherent* (Steel 2013; 2014, Ch. 2; Sunstein, 2005, Ch. 1). Sunstein’s argument can be read in at least two ways (Sandin, 2006, pp. 176-177). On one reading, the charge is that the precautionary principle (in its non-trivial forms) is absolutist, i.e. that it prohibits any policy that has some chance of leading to some harm. Since every policy has a chance of leading to harm, every policy is strictly speaking prohibited. The precautionary principle only *seems* to be action-guiding, according to Sunstein, because various cognitive biases lead people to focus on some risks and ignore others (Sunstein, 2005, Ch. 2). On the other reading, precautionary policies sometimes produce risks that are worse than the ones they were supposed to prevent. In such cases, it would be *inconsistent* to implement the precautionary policy, since it does worse than the alternative according to the very criteria used to justify it.

The absolutist challenge can be solved by setting thresholds of harm and evidence higher than mere logical possibility that any harm will occur. Christian Munthe (2011) suggests that it is a (minimal) requirement for a justifiable version of the precautionary principle that it does not lead to paralysis, which seems to me right. However, this does not solve the inconsistency challenge. Steel solves that challenge by building a requirement of consistency into his account of the precautionary principle. Consistency demands that the precautionary measure suggested “should not be precluded by the same version of [the precautionary principle] used to justify it” (Steel, 2014, p. 28). In other words, the precautionary measures must not itself trigger a demand for precautionary action, using the harm- and evidence thresholds that triggered precautionary action in the first place. Consistency an-

swers Sunstein's challenge (on its second reading), since it limits eligible precautionary actions to those that do not create worse risks than the original activity.

While I agree with Steel that his account of the precautionary principle solves Sunstein's challenges, I do not believe that this will necessarily satisfy those that object to the precautionary principle on the basis of rationality. Sunstein's challenges rely on a fairly strong notion of irrationality, since they suggest that the precautionary principle leads to internally inconsistent or incoherent policies. But rationality objections may be based on weaker notions. Furthermore, Steel's handling of another rationality challenge, put forward by Martin Peterson (2006) seems to me to be similarly based on rejecting unnecessarily strong claims from the critic of precaution. In this case the claim is that certain variations in the likelihoods of outcomes can counterbalance each other *precisely*, which is implausible when we have must rely on purely qualitative information (Steel, 2014, p. 40-41). Several comments made in various seminars by precaution-skeptics have further strengthened my suspicion that Steel's refutations of Sunstein's and Peterson's objections are not entirely satisfactory. As one such philosopher argued, the real challenge to the precautionary principle is that the alternative, namely expected utility maximization, works so well. Article 3 thus defends a Steel-style precautionary principle from rationality objections that are grounded in the ideas underlying expected utility theory – i.e. in the kinds of claims used by proponents of expected utility theory to argue that their approach is *uniquely* rational.

My defense of the precautionary principle mainly concerns the *very use* of thresholds in risk management. As such, it is a defense of the tripod schema, rather than of any particular version of the precautionary principle. Munthe (2015) has criticized Steel for neglecting (or rather refusing) to address the question of when a particular *version* of the precautionary principle is justified, and consequently neglecting to address the ethical justification for precaution. To a certain extent, I sympathize with Steel's approach. The justification of a

given version of the principle will depend on a number of context-specific details, and the question of what level of protection of health and environment is desirable seems to be within the remit of democratic decision-making. At the very least, the rationality-related aspects of justification that Steel addresses are complementary to the purely normative aspects that concern Munthe – they show what kinds of relationships between the involved values are needed if the precautionary principle is to be justified, but are silent on which values bear this relationship to each other.

Nevertheless, my defense of the precautionary schema does invoke ethical theory to a larger extent than Steel. In particular, I suggest that several strands of moral theory justify the use of thresholds of harm and knowledge when possible bad outcomes are disproportionate to possible good outcomes (and that disproportionality occurs more frequently than some might think). Furthermore, I argue that some restrictions on what values *H* and *K* may take are necessary to defend the use of thresholds on purely instrumental terms (that is, defending the use of the precautionary principle even assuming that expected utility maximization is the correct normative theory of choice under uncertainty). Hopefully, my Article thus contributes to the synthesis of instrumentalist and normative arguments that Munthe (2015, p. 221) hopes for.

## **4.2 Cost-benefit analysis**

In Article 3, I suggest that the realistic rival to the precautionary principle is not an ideal version of expected utility theory, but rather a non-ideal version, namely cost-benefit analysis. In that Article, I argue that cost-benefit analysis is not necessarily closer to ideal expected utility theory than the precautionary principle in conditions of uncertainty. The reason is that cost-benefit analysis requires relatively precise estimation of its input parameters, i.e. the utilities and likelihoods of various possible outcomes. This introduces inevitable

estimation errors into the analysis, which results in divergence from ideal expected utility theory, and which are likely to be worse than the errors introduced by the precautionary principle. Steel argues that cost-benefit analysis leads to *paralysis* in cases of uncertainty, since the lack of precise quantitative inputs “no action can be unambiguously justified in its terms” (Steel, 2014, p. 23). My argument is largely complementary to Steel’s; his argument assumes that cost-benefit analysis must or will respect uncertainty, while mine concerns the problems when uncertainty is not respected. But nevertheless Steel’s argument seems to me incomplete. He assumes that if cost-benefit analysis does not unambiguously justify any action, no action will be taken. But whether that is true depends on the institutional embeddings of cost-benefit analysis.

Cass Sunstein, who is the target of my Article 4, provides a detailed analysis of a case in which uncertainty in the input parameters are high. The case concerns the regulation of arsenic in drinking water, and a cost-benefit analysis of the proposed regulation showed monetized benefits ranging from \$13 million to \$3.4 billion (Sunstein, 2002, Ch. 7). Sunstein draws the following “lessons” about the use of cost-benefit analysis in this case:

“Does all this suggest that CBA is, in cases of this sort, unhelpful? It would not be hard to imagine an affirmative answer to that question. A skeptic might conclude that because the range of uncertainty is so large, any number at all could be justified, and the ultimate decision is essentially “political” or based on “values”. This is not exactly wrong, but it should not be taken as a convincing challenge to CBA. An analysis of benefits and costs cannot resolve the ultimate judgment, but it can certainly inform it. Once we understand the potential effects of different arsenic regulations, and see where the uncertainties come from, we are in a much better position to know what to do. Of course the



decision will be a product of “values”; how could it be otherwise? The point is that the values should be identified as such, so that when the government acts its reasons are transparent and explicable” (Sunstein, 2002, p. 178).

If cost-benefit analysis is merely an informational input into decision-making, then Steel is not right to say that it leads to paralysis. Perhaps the associated “values” lead to paralysis, e.g. because they underwrite a presumption against regulation if benefits do not unambiguously exceed costs, but the analysis itself does not.

If cost-benefit analysis is *merely* an informational input, it is hard to see how it could be objectionable. However, there are two problems with Sunstein’s argument here. First, it is not exactly obvious that cost-benefit analysis is *uniquely* in the position to inform decision-making by accounting for all benefits and costs. As I argue in Article 4, the description of the effects of a regulation is only one part of cost-benefit analysis. It also comprises commensuration of values of different effects by monetization, i.e. by assigning a dollar value to each effect. Other regulatory strategies, including the precautionary principle, can unproblematically base themselves on all the best information available. Second, Sunstein does not, when push comes to shove, limit cost-benefit analysis to the status of an informational tool. In fact he discusses a very large number of legal and institutional uses, many of which include wide scope for discretionary decisions on the part of judges and administrators. Consequently it is difficult to ascertain what that overall effect on risk regulation would be if Sunstein got to implement everything he wanted. But there are at least two proposals that suggest that cost-benefit analysis have a more direct effect on policy than merely informing it. Sunstein argues that (i) regulatory agencies must show that the benefits of a proposed regulation justify its costs, and (ii) that presumptive floors and ceilings for the value of a statistical life should be used. In Article 4, I especially target these two proposals.

My argument challenges Sunstein's inventive defence of cost-benefit analysis as a corrective to the cognitive biases that influence our reasoning about risk. Traditionally, cost-benefit analysis is justified by reference to the ideal of economic efficiency, and hence depends on the amalgamation of liberal and consequentialist considerations that justify that ideal. By arguing that cost-benefit analysis functions as a corrective to cognitive biases, Sunstein hopes to secure an "incompletely theorized agreement" on its use from proponents of all normative theories. I argue that the aspects of cost-benefit analysis that Sunstein's proposal (i) and (ii) above promote are not necessary for correcting the most obvious biases, namely those that give rise to false beliefs about the magnitude of risks.

What (i) and especially (ii) *do* promote is "coherent" valuations of risk. The ideal of coherence implies that the value of a statistical life that different regulations imply should not diverge too much from each other. The intuition that it is somehow irrational to care more about one risk than another when they are similar is well known – and it is not uncommon to hear researchers complain that members of the public are irrational to care about the small risks associated with technology when they also ride their bike without a helmet, which is much more dangerous. This is an instance of a similarity argument, and as such my general arguments in Article 2 apply to these as well. In Article 4, I discuss three concrete arguments that Sunstein makes for accepting the coherence ideal, and for believing that cost-benefit analysis promotes coherence in the ways that are supported by these arguments. In each case, I argue, the use of cost-benefit analysis is not supported.

#### **4.3 The regulation of synthetic biology risks**

My arguments in Articles 3 and 4 do not directly address the risks and uncertainties of synthetic biology specifically, but rather address the more general question of what strategy should be used for risk regulation (although I do use examples from the synthetic biology

case in Article 3). Their scope is also limited; I have only shown that the precautionary principle is robust to a certain criticism, and that one defence of cost-benefit analysis fails. Consequently, my contribution to answering question (II) – how synthetic biology should be regulated with respect to risks to health and environment – is likewise limited and general. Nevertheless, I want to finish by saying a little about the conclusions that can be drawn about the regulation of synthetic biology specifically.

Most importantly, the precautionary principle, on the account of it I have defended – and any account that is defensible, as far as I can tell – does not justify a moratorium on “the release and commercial use of synthetic organisms” as the civil society groups demand (FOE et al., 2012, p. 3). Most, and probably all, current commercial applications do not present any meaningful risks to health or the environment – e.g. because they are used only in contained facilities, because they would not be able to survive in the wild, or because they are not remotely likely to be harmful. At best (or worst, depending on your perspective), the precautionary principle justifies pre-release testing of organisms in non-contained use, e.g. as pesticides. Even in such a scheme, there is scope for the use of similarities with other organisms as important evidence for safety. For example, there seems to be little reason to impose strict regulation of yeasts modified to metabolize vanillin (rather than alcohol or carbon dioxide) given that no significant harms has as yet resulted from thousands of years of using of yeast for baking and brewing (except, of course, the harms from the consumption of the metabolite in the latter case). Furthermore, existing regulations already have to deal with the intentional release of microorganisms, and thus should not be unfamiliar with the possible problems posed by the ability of organisms to survive and reproduce.

A more novel problem posed by synthetic biology is the regulation of organisms that are radically different from existing ones. These organisms create problems for risk assessment, since we do not have any data on how they behave in an uncontained environment.

Uncertainty about precise effects will likely be high. But once again, it is hard to construct an argument for a ban or a similarly restrictive policy. For such a ban to be justified on the basis of the precautionary principle, we would have to have some evidence (empirical or theoretical) that a given application may lead to serious damage, and alternative regulations that safeguards against that damage while being less restrictive would have to be lacking. Whether that is the case cannot be settled by philosophical argument.

Similarly, the factual details are important in the case of gain-of-function research, which clearly introduces the possibility of extremely bad outcomes (such as an H5N1 pandemic, which could kill more than 100 million people). If there is some way of regulating the conduct of such research that sufficiently lowers the likelihood of a very bad outcome, such as higher laboratory safety standards and limited permissions to conduct such research, then that is clearly preferable to a ban. Even if that is not possible, the justifiability of a ban depends on the likelihood that the relevant research will result in the *prevention* of a similarly bad outcome. If H5N1 gain-of-function research would enable us to avoid a pandemic, or massively limit the fatalities of such a pandemic, and if it is at least as likely that the research will prevent a pandemic as that it will cause one, then the precautionary principle would not tell against such research.

At some point, synthetic biology is likely to give rise to products with risks of the more ordinary kind – where we can establish a statistical correlation between the product and some adverse effect. If my argument in Article 4 is sound, it would not be irrational not to use cost-benefit analysis as the basis for the regulation of such products. This is a relatively limited conclusion, and cost-benefit analysis may still be warranted on other grounds. But I would suggest (and do suggest in the Article) that the distribution of risks and benefits matter quite a lot, both for the legitimacy of specific regulations and for risk perception. Consider the neighbouring case of GM food. It is often argued that these have massive ben-

efits, and furthermore have no risks. But suppose they posed *some* small risk to the health of those eating them. Since very few of the benefits of GM agriculture accrues to consumers, it seems entirely reasonable for them to demand regulations – not necessarily bans, but at least rules that enable people to avoid GM food. Arguing that the benefits to farmers are larger than the health costs to consumers, and that regulations are therefor unjustified, is not *prima facie* plausible. This does not show that opposition to GM food is warranted, but only that it relies on the fact that the risks are zero. If ordinary people are to be convinced to support technologies that they are not certain are safe, then they need to be convinced that the technology has benefits that they recognize as important.

## References

- Acevedo-Rocha, C.G. (2016). The Synthetic Nature of Biology. In K. Hagen, M. Engelhard & G. Toepfer (Eds.), *Ambivalences of Creating Life: Societal and Philosophical Dimensions of Synthetic Biology*. Cham: Springer.
- Ancilotti, M. & Eriksson, S. (2016). Synthetic Biology in the Press. In K. Hagen, M. Engelhard & G. Toepfer (Eds.), *Ambivalences of Creating Life: Societal and Philosophical Dimensions of Synthetic Biology*. Cham: Springer.
- Attfield, R. (2012). Biocentrism and Artificial Life, *Environmental Values*, 21(1), 83-94.
- Baertschi, B. (2012a). Defeating the Argument from Hubris. *Bioethics*, 27(8), 435-441.
- Baertschi, B. (2012b). The Moral Status of Artificial Life, *Environmental Values*, 21(1), 5-18.
- Bar-Yam, S., Byers-Corbin, J., Casagrande, R., Eichler, F, Lin, A., Oesterreicher, M. ... Oye, K.A. (2012). *The Regulation of Synthetic Biology: A Guide to United States and European Union Regulations, Rules and Guidelines*. Retrieved from [https://www.synberc.org/sites/default/files/Concise%20Guide%20to%20Synbio%20Regulation%20OYE%20Jan%202012\\_0.pdf](https://www.synberc.org/sites/default/files/Concise%20Guide%20to%20Synbio%20Regulation%20OYE%20Jan%202012_0.pdf)
- Basl, J. & Sandler, R. (2013). The good of non-sentient entities: Organisms, artifacts and synthetic biology, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 697-705.

- Bedau, M., Parke, E., Tangen, U. & Hantsche-Tange, B. (2009). Social and ethical checkpoints for bottom-up synthetic biology, or protocells, *Systems and Synthetic Biology*, 3, 65-75.
- Benner, S.A. & Sismour, A.M. (2005). Synthetic Biology, *Nature Reviews Genetics*, 6, 533-543.
- Bergeson, L.L., Campbell, L.M., Dolan, S.L., Engler, R.E., Baron, K.F., Auerbach, B. ... Auer, C.M. (2015). *The DNA of the U.S. Regulatory System: Are We Getting It Right for Synthetic Biology*. Retrieved from [https://www.wilsoncenter.org/sites/default/files/synbio\\_reg\\_report\\_final.pdf](https://www.wilsoncenter.org/sites/default/files/synbio_reg_report_final.pdf)
- Boldt, J. (2013a). Life as a Technological Product: Philosophical and Ethical Aspects of Synthetic Biology. *Biological Theory*, 8, 391-401
- Boldt, J. (2013b). Creating Life: Synthetic Biology and Ethics. In G.E. Kaebnick & T.H. Murray (Eds.), *Synthetic Biology and Morality: Artificial Life and the Bounds of Nature*. Cambridge, MA: MIT Press.
- Boldt, J. (2013c). Do We Have a Moral Obligation to Synthesize Organisms to Increase Biodiversity? On Kinship, Awe and the Value of Life's Diversity, *Bioethics*, 27(8), 411-418.
- Boldt, J., & Müller, O. (2008). Newtons of the Leaves of Grass. *Nature Biotechnology* 26, 387-389.

- Boyle, R.J. & Savulescu, J. (2001). Ethics of using preimplantation genetic diagnosis to select a stem cell donor for an existing person, *British Medical Journal*, 323, 1240-1243.
- Butler, D. (2014). Biosafety controls come under fire, *Nature*, 511, 515-516.
- Callicott, J.B. (1980). Animal Liberation: A Triangular Affair, *Environmental Ethics*, 2, 311-338.
- Caplan, A. (2009). Moving ahead but with greater controls, *Nature Biotechnology*, 27(12), 1105.
- Caplan, A. (2010, September 13). *Testimony to the Presidential Commission for the Study of Bio-ethical Issues*. Retrieved from <http://bioethics.gov/sites/default/files/Testimony-of-Arthur-L-Caplan.pdf>
- Cello, J., Paul, A.V. & Wimmer, E. (2002): Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template, *Science*, 297, 1016-1018.
- Cho, M., Magnus, D., Caplan, A.C., McGee, D. & the Ethics of Genomics Group (1999). Ethical Considerations in Synthesizing a Minimal Genome, *Science*, 286, 2087 + 2089-2090.
- Church, G.M. (2004, August 6). *A Synthetic Biohazard Non-proliferation Proposal*. Retrieved from [http://arep.med.harvard.edu/SBP/Church\\_Biohazard04c.htm](http://arep.med.harvard.edu/SBP/Church_Biohazard04c.htm)



Cohen, G.A. (2011). Rescuing Conservatism. In R.J. Wallace, R. Kumar & S. Freeman (Eds.), *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. Oxford: Oxford University Press.

Deckers, J. (2005): Are Scientists Right and Non-Scientists Wrong? Reflections on Discussions of GM. *Journal of Agricultural and Environmental Ethics*, 18, 451-478

DeLoache, W.C., Russ, Z.N., Narcross, L., Gonzales, A.M., Martin, V.J.J. & Dueber, J.E. (2015). An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose, *Nature Chemical Biology*, 11, 465-471.

Deplazes-Zemp, A. (2012). The moral impact of synthesising living organisms: Biocentric views on synthetic biology, *Environmental Values*, 21(1), 63-82.

Douglas, T., Powell, R., & Savulescu, J. (2013). Is the Creation of Artificial Life Morally Significant? *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 688-696.

Douglas, T. & Savulescu, J. (2010). Synthetic biology and the ethics of knowledge, *Journal of Medical Ethics*, 36, 687-693.

Dormitzer, P.R., Suphaphiphat, P., Gibson, D.G., Wentworth, D.E., Stockwell, T.B., Algire, M.A. ... Venter, J.C. (2013). Synthetic Generation of Influenza Vaccine Viruses for Rapid Response to Pandemics, *Science Translational Medicine*, 5(185ra68), 1-12.

Ehrenberg, R. (2015). Engineered yeast paves way for home-brew heroin, *Nature*, 521, 267-268.

Elliot, R. (1982). Faking Nature, *Inquiry*, 25, 81-93.

European Group on Ethics in Science and New Technology (2009, November 17). *Ethics of Synthetic Biology*. Retrieved from [http://ec.europa.eu/archives/bepa/european-group-ethics/docs/opinion25\\_en.pdf](http://ec.europa.eu/archives/bepa/european-group-ethics/docs/opinion25_en.pdf)

Endy, D. (2005). Foundations for Engineering Biology. *Nature*, 438, 449-453.

ETC Group (2014, June 2). *Open Letter to Ecover/Method*. Retrieved from <http://www.etcgroup.org/content/open-letter-ecover-method>

European Commission (2000, February 2). *Communication from the Commission on the Precautionary Principle*. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52000DC0001&from=EN>

European Commission (2014, September 23). *Opinion on Synthetic Biology I: Definition*. Retrieved from [http://ec.europa.eu/health/scientific\\_committees/emerging/docs/scenihr\\_o\\_044.pdf](http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_044.pdf)

European Commission (2015a, April 29). *Opinion on Synthetic Biology II: Risk assessment methodologies and safety aspects*. Retrieved from [http://ec.europa.eu/health/scientific\\_committees/emerging/docs/scenihr\\_o\\_048.pdf](http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_048.pdf)

European Commission (2015b, November 27). *Opinion on Synthetic Biology I: Risks to the environment and biodiversity*. Retrieved from [http://ec.europa.eu/health/scientific\\_committees/emerging/docs/scenihr\\_o\\_050.pdf](http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_050.pdf)

FOE, ICTA & the ETC Group (2012, March 13). *Principles for the Oversight of Synthetic Biology*. Retrieved from [http://libcloud.s3.amazonaws.com/93/ae/9/2287/2/Principles\\_for\\_the\\_oversight\\_of\\_synthetic\\_biology.pdf](http://libcloud.s3.amazonaws.com/93/ae/9/2287/2/Principles_for_the_oversight_of_synthetic_biology.pdf)

French, C.E., de Mora, K., Joshi, N., Elfick, A., Haseloff, J. & Ajioka, J. (2011). Synthetic Biology and the Art of Biosensor Design. In E.R. Choffnes, D.A. Relman & L. Pray (Eds.), *The Science and Applications of Synthetic and Systems Biology*. Washington, DC: The National Academies Press.

Garfinkel, M.S., Endy, D., Epstein, G.L. & Friedman, R.M. (2007). *Synthetic Genomics: Options for Governance*. Retrieved from <http://www.jcvi.org/cms/fileadmin/site/research/projects/synthetic-genomics-report/synthetic-genomics-report.pdf>

Garnett, K. & Parsons, D.J. (2016). Multi-Case Review of the Application of the Precautionary Principle in European Union Law and Case Law, *Risk Analysis*, doi:10.1111/risa.12633

Gaskell, G., Stares, S., Allansdottir, A., Allum, N., Castro, P., Esmer, Y., ... Wagner, W. (2010). Europeans and biotechnology in 2010: Winds of change? *Report to the European Commissions Directorate-General for Research*. Retrieved from [http://ec.europa.eu/public\\_opinion/archives/ebs/ebs\\_341\\_winds\\_en.pdf](http://ec.europa.eu/public_opinion/archives/ebs/ebs_341_winds_en.pdf).

- Gaus, G. (2005). The Place of Autonomy within Liberalism. In J. Christman & J. Anderson (Eds.), *Autonomy and the Challenges to Liberalism*. Cambridge: Cambridge University Press.
- Gelanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F. & Smolke, C.D. (2015). Complete biosynthesis of opioids in yeast, *Science*, 349, 1095-1100.
- Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.-Y., Algire, M.A. ... Venter, J.C. (2010). Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome, *Science*, 329, 52-56.
- Hart Research Associates (2013, March 6). *Awareness & Impressions Of Synthetic Biology*. Report to the Synthetic Biology Project of the Woodrow Wilson International Center for Scholars. Retrieved from <https://www.wilsoncenter.org/publication/awareness-impressions-synthetic-biology>.
- Hart Research Associates (2014, April 18). *Perceptions of Synthetic Biology and Neural Engineering*. Report to the Synthetic Biology Project of the Woodrow Wilson International Center for Scholars. Retrieved from [http://www.synbioproject.org/site/assets/files/1293/focusgroup\\_2014.pdf](http://www.synbioproject.org/site/assets/files/1293/focusgroup_2014.pdf)
- Heath, J. (2011). Three Normative Models of the Welfare State, *Public Reason*, 3(2), 13-43.
- Heavey, P., (2013). Synthetic Biology Ethics: A Deontological Assessment, *Bioethics*, 27(8), 442-452.

- Heyd, D. (2012). Is there anything unique in the ethics of synthetic biology? *Perspectives in Biology and Medicine*, 55(4), 55:581–589.
- Holm, S. (2012a). The Scientific Aspirations of Synthetic Biology and the Need for Analytic Ethics. *Ethics, Policy & Environment*, 15(1), 25-28.
- Holm, S. (2012b). Biological Interests, Normative Functions, and Synthetic Biology, *Philosophy and Technology*, 25, 525-541.
- Huesken, S. (2014). Artificial Life and Ethics, *Nanoethics*, 8,111-116
- Hutchison, C.A., Chuang, R.-Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H. ... Venter, J.C. (2016). Design and synthesis of a minimal bacterial genome, *Science*, 351, 1414.
- Jefferson, C., Lentzos, F. & Marris, C. (2014). Synthetic biology and biosecurity: challenging the myths, *Frontiers in Public Health*, 2, 1-15.
- Jermey, A. (2011). Synthetic biology: Licensing bacteria to kill, *Nature Reviews Microbiology*, 9, 699.
- Jiang, P. (2014). A Uniform Precautionary Principle Under EU Law, *PKU Transnational Law Review*, 2(2), 491-518.

- Johnson, N. (2014, June 5). The next front in the GMO war: Synthetic biology. *Grist*. Retrieved from <http://grist.org/business-technology/the-next-front-in-the-gmo-war-synthetic-biology/>
- Johnson, N.P.A.S. & Mueller, J. (2002). Updating the Accounts: Global Mortality of the 1918-1920 "Spanish" Influenza Pandemic, *Bulletin of the History of Medicine*, 76(1), 105-115.
- Kaebnick, G.E. (2009). Should moral objections to synthetic biology affect public policy? *Nature Biotechnology*, 27(12), 1106-1108.
- Kaebnick, G.E. (2012, March 2012). *Carefully Precautionary about Synthetic Biology*. [Web log post]. Retrieved from <http://www.thehastingscenter.org/carefully-precautionary-about-synthetic-biology/#ixzz33iPzJJ5c>
- Kaebnick, G.E. & Murray, T.H. (2013). Introduction. In G.E. Kaebnick & T.H. Murray (Eds.), *Synthetic Biology and Morality: Artificial Life and the Bounds of Nature*. Cambridge, MA: MIT Press.
- Kaiser, J. (2014, May 27). U.S. advisers sign of on plan for reviewing risky virus studies, *ScienceInsider*. Retrieved from <http://www.sciencemag.org/news/2016/05/us-advisers-sign-plan-reviewing-risky-virus-studies>
- Katz, E. (1992). The Big Lie: Human Restoration of Nature, *Research in Philosophy and Technology*, 12, 231-241.

- Keasling, J. (2010). Manufacturing Molecules Through Metabolic Engineering, *Science*, 330, 1355-1358.
- Khalil, A.S. & Collins, J.J. (2010). Synthetic biology: applications come of age, *Nature Reviews Genetics*, 11, 367-379.
- Kronberger, N. (2012): Synthetic biology: taking a look at a field in the making, *Public Understanding of Science*, 21, 130-133.
- Kysar, D.A. (2004). Preferences for Processes: The Process/Product Distinction and the Regulation of Consumer Choice, *Harvard Law Review*, 118(2), 525-642.
- Landes, X. & Holtug, N. (2015). Insurance, Equality and the Welfare State: Political Philosophy and (of) Public Insurance, *Res Publica*, 21, 111-118.
- Lee, K. (1999): *The Natural and the Artefactual: The Implications of Deep Science and Deep Technology for Environmental Philosophy*. Lanham, MD: Lexington Books.
- Lewens, T. (2013). From *bricolage* to BioBricks™: Synthetic biology and rational design, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 641-648.
- Lewidow, L., Carr, S. & Wield, D. (2005). European regulation of agri-biotechnology: precautionary links between science, expertise and policy, *Science and Public Policy*, 32(4), 261-276.

- Lewontin, R.C. (2014, May 8). The New Synthetic Biology: Who Gains? *New York Review of Books*. Retrieved from <http://www.nybooks.com/articles/2014/05/08/new-synthetic-biology-who-gains/>
- Link, H.-J. (2013): Playing God and the Intrinsic Value of Life: Moral Problems for Synthetic Biology? *Science and Engineering Ethics*, 19, 435-448.
- Lo, Y.S. (2010). The Land Ethic and Callicott's Ethical System (1980-2001): An Overview and Critique. *Inquiry: An Interdisciplinary Journal of Philosophy*, 44(3), 331-358.
- Löfstedt, R. (2014). The precautionary principle in the EU: Why a formal review is long overdue, *Risk Management*, 16(3), 137-163.
- McGarity, T.O. (2002). Seeds of Distrust: Federal Regulation of Genetically Modified Foods, *University of Michigan Law Review*, 35(3), 403-510.
- Miller, S. & Selgelid, M.J. (2007). Ethical and Philosophical Consideration of the Dual-Use Dilemma in the Biological Sciences, *Science and Engineering Ethics*, 13, 523-580.
- Munthe, C. (2011). *The Price of Precaution and the Ethics of Risk*. Dordrecht: Springer.
- Munthe, C. (2015). Precaution, bioethics and normative justification, *Monash Bioethics Review*, 33, 219-225.



- Møller, B.L. (2014). Disruptive innovation: Channeling photosynthetic electron flow into light-driven synthesis of high-value products, *Synthetic Biology*, 1, 330-359.
- National Research Council (2004). *Biotechnology Research in an Age of Terrorism*. Retrieved from <http://www.nap.edu/catalog/10827/biotechnology-research-in-an-age-of-terrorism>
- NSABB (2012, March 29-30). *Findings and Recommendations*. Retrieved from [https://www.nih.gov/sites/default/files/about-nih/nih-director/statements/collins/03302012\\_NSABB\\_Recommendations.pdf](https://www.nih.gov/sites/default/files/about-nih/nih-director/statements/collins/03302012_NSABB_Recommendations.pdf)
- Nuffield Council on Bioethics (2012, December). *Emerging biotechnologies: Technology, choice and the public good*. Retrieved from <http://nuffieldbioethics.org/project/emerging-biotechnologies/>
- O'Malley, M., Powell, A., Davies, J.F., & Calvert, J. (2008) Knowledge-Making Distinctions in Synthetic Biology. *BioEssays*, 30(1), 57-65.
- Osterholm, M.T. & Kelley, N.S. (2012). Mammalian-Transmissible H5N1 Influenza: Facts and Perspective, *mBio*, 3(2), 1-4.
- Pamlin, D. & Armstrong, S. (2015). *12 Risks that threaten human civilization: The case for a new risk category*. Global Challenges Foundation. Retrieved from <http://globalchallenges.org/wp-content/uploads/12-Risks-with-infinite-impact.pdf>

Parens, E., Johnston, J. & Moses, J. (2008). Do We Need "Synthetic Bioethics"? *Science*, 321, 1449.

Parens, E., Johnston, J. & Moses, J. (2009, June). *Ethical Issues in Synthetic Biology*. Retrieved from <http://www.synbioproject.org/publications/synbio3/>

Pauwels, E. (2009): Review of quantitative and qualitative studies on U.S. public perceptions of synthetic biology. *Systems and Synthetic Biology*, 3, 37-46.

Perkel, J.M. (2015, June 9). Revolutionizing Vaccine Development with Synthetic Biology. *Biocompare*. Retrieved from <http://www.biocompare.com/Editorial-Articles/174945-Revolutionizing-Vaccine-Development-with-Synthetic-Biology/>

Pierce, R.L. (2012). Whose ethics of knowledge? Taking the next step in evaluating knowledge in synthetic biology: a response to Douglas and Savulescu, *Journal of Medical Ethics*, 38, 636-638.

PCSBI (2010). *New Directions: The Ethics of Synthetic Biology and Emerging Technologies*. Washington, DC: Presidential Commission for the Study of Bioethical Issues.

Peterson, M. (2003). Transformative Decision Rules, *Erkenntnis*, 58(1), 71-85.

Peterson, M. (2006). The Precautionary Principle is Incoherent, *Risk Analysis*, 26(3), 595-601.

Pollack, A. (2002, July 12). Scientists Create a Live Polio Virus. *New York Times*. Retrieved from <http://www.nytimes.com/2002/07/12/science/12POLI.html?pagewanted=all>

Pollack, A., (2013, October 20). What's That Smell? Exotic Scents Made From Re-engineered Yeast. *New York Times*. Retrieved from [http://www.nytimes.com/2013/10/21/business/whats-that-smell-exotic-scents-made-from-re-engineered-yeast.html?\\_r=1](http://www.nytimes.com/2013/10/21/business/whats-that-smell-exotic-scents-made-from-re-engineered-yeast.html?_r=1)

Powell, R. (2015). Adopting a Technological Stance Toward the Living World. Promises, Pitfalls and Perils. In S.O. Hansson (Ed.), *The Role of Technology in Science*. Dordrecht: Springer.

Preston, B. (2013). Synthetic Biology as Red Herring. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 649-659.

Preston, C.J. (2008). Synthetic Biology: Drawing a Line in Darwin's Sand. *Environmental Values*, 17(1), 23-39

Preston, C.J. (2013a). Synthetic Bacteria, Natural Processes and Intrinsic Value. In G.E. Kaebnick & T.H. Murray (Eds.), *Synthetic Biology and Morality: Artificial Life and the Bounds of Nature*. Cambridge, MA: MIT Press.

Preston, C.J. (2013b). Evolution and the Deep Past: Intrinsic Responses to Synthetic Biology. In R.L. Sandler (Ed.), *Ethics and Emerging Technologies*. Basingstoke: Palgrave Macmillan.

Randall, A. (2011). *Risk and Precaution*. Cambridge: Cambridge University Press.

Rappert, B. & Selgelid, M.J. (Eds.) (2013). *On the Dual Uses of Science and Ethics: Principles, Practices and Prospects*. Canberra: ANU Press.

Reardon, S. (2014). US suspends risky disease research, *Nature*, 514, 411-412.

Ribeiro, Silvia & Thomas, Jim (2015, October 23): A Bad Bet on Synthetic Biology. *Project Syndicate*. Retrieved from <https://www.project-syndicate.org/commentary/stevia-synthetic-biology-dangers-by-silvia-ribeiro-and-jim-thomas-2015-10>.

Rodemeyer, M. (2009). *New Life, Old Bottles: Regulating First-Generation Products of Synthetic Biology*. Retrieved from [http://www.synbioproject.org/process/assets/files/6319/nano\\_synbio2\\_electronic\\_final.pdf](http://www.synbioproject.org/process/assets/files/6319/nano_synbio2_electronic_final.pdf)

Sandin, P. (1999). Dimensions of the Precautionary Principle, *Human and Ecological Risk Assessment*, 5(5), 889-907.

Sandin, P., Peterson, M., Hansson, S.O., Rudén, C. & Juthe, A. (2002). Five charges against the precautionary principle, *Journal of Risk Research*, 5(4), 287-299.

Sandler, R. (2004). An Aretaic Objection to Agricultural Biotechnology, *Journal of Agricultural and Environmental Ethics*, 17, 301-317.

Sandler, R. (2007). *Character and Environment*. New York: Columbia University Press.

- Schmidt, M., Ganguli-Mitra, A., Torgersen, H., Kelle, A., Deplazes, A. & Biller-Andorno, N. (2009). A priority paper for the societal and ethical aspects of synthetic biology, *Systems and Synthetic Biology*, 3, 3-7.
- Scott, D. (2011). The Technological Fix Criticisms and the Agricultural Biotechnology Debate, *Journal of Agricultural and Environmental Ethics*, 24, 207-226.
- Selgelid, M.J. (2007). A Tale of Two Studies: Ethics, Bioterrorism and the Censorship of Science, *Hastings Center Report*, 37(3), 35-43.
- Selgelid, M.J. (2016). Gain-of-Function Research: Ethical Analysis, *Science and Engineering Ethics*, doi:10.1007/s11948-016-9810-1
- Service, R.F. (2015, May 18). Researchers closer to engineering yeast that make morphine, spurring worries. *ScienceInsider*. Retrieved from <http://www.sciencemag.org/news/2015/05/researchers-closer-engineering-yeast-make-morphine-spurring-worries>
- Slomovic, S., Pardee, K. & Collins, J.J. (2015). Synthetic biology devices for in vitro and in vivo diagnostics, *PNAS*, 112(47), 14429-14435.
- Smith, K. (2013). Synthetic Biology: A Utilitarian Perspective, *Bioethics*, 27(8), 453-463.
- Steel, D. (2013). The Precautionary Principle and the Dilemma Objection, *Ethics, Policy and Environment*, 16(3), 321-340.

- Steel, D. (2014). *Philosophy and the Precautionary Principle: Science, Evidence and Environmental Policy*. Cambridge: Cambridge University Press.
- Steurer, W. (2016). "Some Kind of Genetic Engineering... Only One Step Further"—Public Perceptions of Synthetic Biology in Austria. In K. Hagen, M. Engelhard & G. Toepfer (Eds.), *Ambivalences of Creating Life: Societal and Philosophical Dimensions of Synthetic Biology*. Cham: Springer.
- Stiglitz, J.E. & Greenwald, B.N. (2014): *Creating a Learning Society*. New York: Columbia University Press.
- Strawson, P.F. (1962). Freedom and Resentment, *Proceedings of the British Academy*, 48, 1-25.
- Sunstein, C.R. (2002). *Risk and Reason: Safety, Law and the Environment*. Cambridge: Cambridge University Press.
- Sunstein, C.R. (2005). *Laws of Fear: Beyond the Precautionary Principle*. Cambridge: Cambridge University Press.
- Thomas, Jim (2013, April 12): Synthetic anti-malarial compound is bad news for artemisia farmers. *The Guardian*. Retrieved from <http://www.theguardian.com/global-development/poverty-matters/2013/apr/12/synthetic-malaria-compound-artemisia-farmers>
- Thompson, P.B. (2012). Synthetic Biology Needs A Synthetic Bioethics, *Ethics, Policy and Environment*, 15(1), 1-20.

- Thompson, P.B. (2015). From Synthetic Bioethics to One Bioethics: A Reply to Critics, *Ethics, Policy and Environment*, 18(2), 215-224.
- Trouwborst, A. (2006). *Precautionary Rights and Duties of States*. Leiden: Martinus Nijhoff.
- Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solórzano, A., Swayne, D.E. ... Garcia-Sastre, A. (2005). Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus, *Science*, 310, 77-80.
- Von Bubnoff, A. (2005). The 1918 flu virus is resurrected, *Nature*, 437, 794-795.
- Watson, E. (2014, August 29). So Häagen-Dazs 'says no to synbio'? *Food Navigator*. Retrieved from <http://www.foodnavigator-usa.com/Manufacturers/So-Haeagen-Dazs-says-no-to-synbio>
- Wimmer, E., Mueller, S., Tumpey, T.M. & Taubenberger, J.K. (2009). Synthetic viruses: A new opportunity to understand and prevent viral disease, *Nature Biotechnology*, 27, 1163-1172.
- Yang, S. (2015, May 18). Discovery paves way for homebrewed drugs, prompts call for regulation. *Berkeley News*. Retrieved from <http://news.berkeley.edu/2015/05/18/brewing-opioids-scientists-urge-regulation/>





# Article 1 Synthetic Biology and the Moral Significance of Artificial Life

## **1. Introduction**

In 2010, the J. Craig Venter Institute (JCVI) announced that they had created the first cell with a wholly synthetic genome. This was a milestone in the development of synthetic biology, and was widely publicised as an important step towards creating artificial life. It spurred much debate, including debate concerning the ethical implications of creating artificial life and of synthetic biology generally. An important strand of that discussion concerned whether these developments were, in fact, morally significant. In a 2013 paper, Thomas Douglas, Russell Powell and Julian Savulescu (henceforth DPS) argue that creating artificial life is not morally significant.

In this paper, I critique their argument. I argue that the definitions of ‘artificial life’ and of ‘moral significance’ are unnecessarily narrow: They do not facilitate the best interpretations of ethical objections to creating artificial life, and they are not well motivated on independent grounds. In the first two sections, I examine the notions, starting with artificial life, and following with moral significance. In each case, I aim to show that DPS’ conceptions do not uniquely capture the sense of the relevant concept, and that they are not well suited to the subject matter, namely synthetic biology. In the final section, I attempt to show how important critical views on synthetic biology are more plausible when interpreted in line with my more expansive understanding of artificial life and of moral significance. Therefore, I conclude, DPS have not succeeded in showing that creating artificial life, considered as an aspect of synthetic biology, is not morally significant.

## **2. Creation of artificial life**

In this section, I discuss the way DPS understand the activity that they are evaluating. In this section, I argue that the definition of ‘artificial life’ that DPS rely on is poorly suited for evaluating the creation of artificial life *as that concept applies to synthetic biology*. The definition neither tracks the core projects of synthetic biology nor the ethical concerns about it. Before making that argument, I will argue that there are several equally plausible conceptions of artificial life, of which DPS discuss only one.

### **2.1 Artificial life**

DPS’ definition of artificial life is based on the assumption that “[t]he creation of artificial life would have to consist ... in the creation of an artificial *living entity*, henceforth an ‘artificial organism’” (Douglas et al., 2013, p. 689). They then go on to offer the following definition of when an organism is artificial:

We will take an organism to be artificial just in case *either* (1) all core elements of that organism were initially constructed from chemically simple, non-living materials to the specification of a person or other natural rational being, or (2) it descended from an organism (or pair of organisms in the case of sexual creatures) that was constructed in this way. (Douglas et al. 2013, p. 689)

This definition involves a certain conception of artificiality (as that concept is applied to life), concerning the *subject* of artificiality – or in other words what phenomenon artificiality is predicated of. For any conception of artificial life, the subject will have to be *life* or *living things*. But these notions have several meanings, yielding several possible conceptions of artificiality; each corresponding to a sense of ‘life’ or ‘living thing’ of which artificiality is predicated. We can distinguish between (at least) three conceptions.

*Token artificiality.* This is the conception that is assumed by DPS. It sees the subject of artificiality as the individual (token) organism.

*Type artificiality.* On the second conception, the subject of artificiality is not the individual organism but the type or kind of organism. The description ‘artificial life form’ may be more apt than ‘artificial organism’ for this conception. On this conception, genetically modified organisms (GMOs) are artificial, and so, arguably, are domesticated and selectively bred plants and animals (Preston 2013; Sperber 2007).

*State-of-being artificiality.* Finally, the subject of artificiality could be the very livingness of living things. Here, it is those definitive features of living things in virtue of which they count as living that are recreated in artificial form. This would include such things as homeostasis, growth and reproduction. Examples include organisms not based on DNA, and (on some views) computer simulations of life. The term ‘artificial life’ best applies to this.

I have only described the various conceptions here in a very rough-and-ready form. My suggestion is that neither corresponds more precisely to our linguistic intuitions with respect to ‘artificial life’ than the others.

The JCVI’s 2010 announcement sparked considerable debate concerning whether or not artificial life had been achieved. DPS note this and refrain from taking a stance on the question. But they do take a stance on what is at issue in that debate. On the claim that the JCVI had created artificial life is “doubtful ... given that only the genome and none of the cytoplasmic structures were synthesized by scientists” (Douglas et al., 2013, p. 688). The implication is that if the cytoplasmic structures had also been synthetic, artificial life would

have been created. But this point was also disputed, on the grounds that such a bacterium would merely be a copy of a naturally occurring life form (Bedau et al., 2010; see especially the comments by George Church, Steen Rasmussen, Martin Fussenegger and Jim Collins). For example, Steen Rasmussen argues that creating artificial life would require “constructing life using different materials and blueprints” (Bedau et al., 2010, p. 423), i.e. would require a state-of-being artificial organism. On the other side, Martin Fussenegger argues, the JCVI bacterium was “a technical advance, not a conceptual one”, and that “[t]his latest technology will simply increase the speed with which new organisms can be generated” (Bedau et al., 2010, p. 424), implying that type artificiality is the relevant conception.

It may be that one conception is superior to another in some respect (e.g. for learning about the nature of life or disproving vitalism). My claim here is the modest one that none of the three conceptions better corresponds to the ‘folk’ concept of artificial life, simply because there is no unified concept (Machery, 2012). So there is no reason to restrict our understanding of artificial life to just one of the conceptions based on conceptual analysis or linguistic intuitions.

## **2.2 Artificiality and synthetic biology**

In this section, I argue that the DPS definition is ill suited for capturing what is going on within synthetic biology. In order to do so, it will be useful to look more closely at the definition of artificiality.

On the DPS definition, an organism is artificial just in case “all core elements of that organism were initially constructed from chemically simple, non-living materials to the specification of a person” (or the organism was descended from such organisms). We can distinguish two aspects of artificiality in this definition (as DPS themselves do):

*Fabrication.* The fact that “all core elements ... were initially constructed from chemically simple, non-living materials”. We can distinguish between fabrication in general – the bare fact that something was produced, in some sense, by human beings – from the specifics of DPS’ definition of fabrication. For the former, I will use the word ‘production’ and its derivative, reserving ‘fabrication’ for DPS’ criterion.

*Design.* The fact that the organism is created “to the specification of a person”.

DPS’ place more emphasis on fabrication than on design. Fabrication is treated as a necessary condition, since the fact that organisms produced by predecessor technologies (e.g. GMOs) are claimed not to be artificial “since these are not constructed from chemically simple, non-living materials” (Douglas et al., 2013, 689). On the other hand, design seems to be non-necessary, since an organism that is “genetically and phenotypically identical to [a] wild type” can still be artificial (Douglas et al., 2013, 689). One might argue that even a copy is *designed to be* a copy, but in that case the design conditions seems to be automatically met whenever the fabrication condition is. I suggest that DPS’ focus on token artificiality is the cause of this inequality in emphasis. If we want to capture type artificiality, design *is* a necessary condition; fabricated copies of existing life forms are precisely not new life forms.

I argued above that there is no reason to restrict the artificial life to token artificiality on conceptual grounds. My claim now is that the related emphasis on fabrication is not a good reflection of what synthetic biologists are doing – synthetic biologists take themselves to be designing new life forms just as much as fabricating individual artificial organisms. DPS focus almost exclusively on the JCVI’s synthetic cell, but synthetic biology is not co-extensional with the project of the JCVI, or with projects like it. On the contrary, synthetic

biology is a broad field, encompassing several different types of project (some might even argue that it is not a unified field at all).

Maureen O'Malley, Alexander Powell, Jonathan F. Davies and Jane Calvert (2008) suggest a rough division of synthetic biology research into three groups: (i) DNA-based device construction, (ii) genome driven cell engineering and (iii) protocell creation. Group (i) is the group that is most directly engaged in developing biologically based technology, and in bringing engineering principles into biology (as outlined in Endy, 2005). One of the poster children for synthetic biology, the modification of *e. coli* and baker's yeast to produce a precursor to the malaria drug artemisinin, also belongs in this group. Group (ii) encompasses the JCVI's project. Its aims include synthesizing entire genomes and producing a minimal cell by isolating the genes required for maintaining basic life processes. Group (iii) includes those researchers who are working to construct "minimal cellular systems" and to understand "fundamental biological properties" by recreating them. O'Malley et al. include *inter alia* George Church's work in this group (O'Malley et al., 2008, p. 59).

There is no reason to think that any one of these groups is more definitive of synthetic biology than the others. Each contains paradigmatic projects and prominent spokespersons for the field. Furthermore, the technological aims of synthetic biology are present in all groups.<sup>1</sup> At the most basic level, those aims consist in creating organisms that have properties that are useful for human beings. Design and type artificiality are clearly essential to such aims – it is the prospect of being better able to create novel life *forms* that suit our desires that is the promise of synthetic biology viewed as a technological endeavour. On the other hand, not all of the projects would place much weight on the fabrication aspect. With-

---

<sup>1</sup> For a discussion of the distinction between scientific and technological aspirations of synthetic biology see (Holm, 2012).

in group (i), for example, the organisms that are created are typically modified versions of existing organisms, such as bacteria, yeasts or algae.

So the field of synthetic biology does not share the emphasis on fabrication that DPS exemplify. If anything, it seems that design and type artificiality is more central. Understanding the ethical implications of artificial life *as an aspect of synthetic biology* thus requires taking design and type artificiality seriously.

### **2.3 Artificiality and ethical concerns**

Even if synthetic biology's conception of artificial life is not restricted to fabrication and token artificiality, it could still be true that ethical objections are grounded in these conceptions. DPS argue that they are. They note that designed organisms have existed for a long time, arguably since the first plants and animals were domesticated more than 10.000 years ago. Given this fact, they argue, "it would be surprising if those alarmed by the prospect of creating artificial life were alarmed by the [design] aspect". Furthermore, if design was what worried people, "it would be difficult to explain why the JCVI's creation was singled out for attention" (Douglas et al., 2013, 689). So DPS pose two explanatory challenges for those who see the design aspect of artificial organisms as a possible basis for ethical objections: (i) Why did the JCVI's specific project receive so much attention, when design played no role in it? (ii) Why does synthetic biology attract ethical attention at all, when designed organisms are also created by many other technologies, from domestication to genetic engineering?

Regarding (i), I am less convinced than DPS that the JCVI was the object of attention, at least in the sense that critics' arguments were aimed directly at what the JCVI had done (or at the culmination of that project). Certainly not all critics concerned themselves much with the details of what the JCVI had done. Even direct reactions to the JCVI, such as the

report of the Presidential Commission for the Study of Bioethical Issues (PCSB I) did not focus exclusively on that project (Presidential Commission for the Study of Bioethical Issues, 2010). Of course, the JCVI's announcement *occasioned* much debate, but this simple fact is not hard to explain; the event was widely publicised, and often described "in ways more provocative than accurate" (PCSB I, 2010, 155). And for philosophers, the announcement presented an opportunity to consider and debate the merits of synthetic biology more broadly.

Now consider (ii), the idea that it is hard to explain why synthetic biology should receive any attention if critics were worried about the design aspect and not about fabrication. Note that the very existence of an explanation problem requires that critics accept the earlier instances of designed organisms as morally unproblematic. While this is plausible for domestication and probably selective breeding, the immediate predecessor to synthetic biology – traditional genetic engineering – has not exactly met universal approval. So the datum to be explained seems more often to be why people should be sceptical of genetic engineering *and* synthetic biology when they accept domestication and (probably) selective breeding. But of course there are critics who, like Joachim Boldt and Oliver Müller, argue that "certain ethical implications of synthetic biology go beyond those of genetic engineering" (Boldt & Müller, 2008). Either way, it seems to be true that there are instances of designed organisms that critics view as unproblematic. So the question remains: How to explain this fact if the basis of ethical misgivings is the fact that synthetic organisms are *designed*?

There are, I think, two main ways in which an explanation could be given. First, one could argue that design only provides grounds for an objection when certain other factors are (or are not) present. The method of production could be one such factor – so, for example, a critic might hold that designing new (type artificial) organisms by way of selecting the most useful specimens for breeding or planting is unproblematic, while doing so by way of



directly changing the organism's genome is troubling. In such cases, design is still a non-redundant part of the basis for objections.

Second, one could argue that differences in *degree* are crucial. For such a view, the range and specificity of human control over the organism's properties are ethically important. This, I think, provides the best interpretation of some prominent examples of critics that stress that synthetic biology is more problematic than genetic engineering. This includes Boldt and Müller, who DPS cite as an example of this type of view. It is true that Boldt and Müller argue that 'creation' is more troubling than 'manipulation'. But at least one of the reasons why this is the case is that created organisms exhibit a more radical kind of design. Whereas manipulation (i.e. genetic engineering) "can be described as softening the unpleasant edges" of an existing organism, creation (i.e. synthetic biology) "does not add value to an existing organism; it brings into existence something that counts as valuable from our point of view". And the core ethical problem seems to be that this entails the view that "nature is a blank space to be filled with whatever we wish" (Boldt & Müller, 2008, p. 388). Furthermore, Boldt and Müller suggest that "the creation of an organism that does not differ from a natural counterpart in any relevant respect" is *less* troubling than standard genetic engineering (Boldt & Müller, 2008, p. 388). In other words, fabrication in itself is not the basis for Boldt and Müller's ethical objections. In §4, I will argue that the same goes for several other prominent arguments.

## **2.4 Taking stock**

I have argued that DPS' restriction of 'artificial' life to token artificiality and fabrication is not warranted. Neither (i) considerations about the proper meaning of the concept of artificial life, (ii) a desire to capture the core of synthetic biology, or (iii) the need to explain why synthetic biology has been the object of ethical objections are sufficient grounds for such a

restriction. I therefore suggest that an examination of the moral significance of creating artificial life should have a broader focus than DPS' paper does – it should focus on the kinds of things synthetic biologists are doing with organisms.

### **3. Moral significance**

In this section, I consider the second facet of the moral significance of creating artificial life, namely *moral significance*. I first consider generally what moral significance might be taken to be, and then discuss DPS' definition of it. I will argue that this definition is too narrow, especially because it ignores the possibility that creating artificial life might be conditionally (but not unconditionally) acceptable.

#### **3.1 Senses of significance**

As was the case for artificial life, there is no single agreed-upon definition of moral significance. And like artificial life, it is useful to distinguish between senses based on what the *subject* of moral significance is. As far as I can tell, there are at least four possibilities:

(B) *Significance of Beings*. Moral significance is an extension of moral consideration. A being is morally considerable just in case we ought to take it (and its interests) into account in our moral deliberations. The moral significance of a considerable being is the weight we ought to give that being (and its interest) in moral deliberation (Goodpaster, 1978, p. 322-323).

(P) *Significance of Properties of Beings*. A property of a being is morally significant if that property plays a part in determining the moral status of the being. Examples of properties that have been argued to be significant include being a person, being sentient and being alive.

(F) *Significance of Features of Actions*. A feature of an action is morally significant if it plays a role in determining the overall moral status of the action. Morally significant features provide reasons for and against performing actions.

(T) *Significance of Action Types*. An action type (or a practice) is morally significant if some moral judgment is true of the type as such. That is, significance is predicated of the whole type rather than individual instances of the type. Only if some moral judgment is true of that whole is the action type significant.

Both (B) and (P) may play a role in the ethics of synthetic biology. For (B) the question would be how much weight to assign to the interests of artificial organisms. (P) is an issue if the discussion concerns whether *artificiality* for organisms ought to alter their moral status. But that does not seem to be what concerns DPS, since their discussion clearly concerns the moral significance of an action or a practice – namely of creating artificial life.

That leaves (F) and (T) as possible sense of moral significance in this case. Since *creating artificial life* is clearly an action type, we might think that (T) is the obvious choice. But that fact in itself does not rule out (F), since *being a token of a type* is also a feature of an action. (F) would ground the claim that the bare fact that an action constitutes the creation of artificial life is a reason not to perform that action. In other words, the property of being an instance of creating artificial life would be a wrong-making feature of any action that has it. This does not mean that all such actions are wrong all things considered. Creating artificial life would more plausibly be a *pro tanto* or contributory wrong-making feature, i.e. a feature that could be outweighed by other, right-making features.

If (T) is the right sense of moral significance, the question is whether some moral judgment attaches to creating artificial life, considered as a type of action. It need not be the case that the relevant moral judgment applies to *each* instance in the set, since some moral

judgments are not apt to be applied to individual instances. I am thinking especially of the judgment that an action type is *conditionally acceptable* – i.e. the judgment actions of the type are permissible under some conditions, but not under others. As I will argue below, this is exactly the judgment that critics of synthetic biology most plausibly make (or, for some, ought to make), and it is a judgment that DPS are unable to capture.

### 3.2 Moral significance and artificial life

Let me now turn from general considerations about moral significance to DPS' specific definition of the concept for the case of creating artificial life. Their definition is this:

We will take [the creation of artificial life] to have such significance just in case

(a) there are moral reasons not to create artificial organisms, or factors that weaken our moral reasons *to* create them, and (b) these are specific to the creation of artificial organisms (Douglas et al., 2013, p. 689).

I have already argued that condition (b) is problematic, since worries based on design would apply to the contrast practices that DPS have in mind, such as genetic engineering. I will not say more about it here.

Consider how condition (a) relates to (F) and (T), the senses of moral significance that are relevant for creating artificial life. If (a) is interpreted in terms of (F), there would have to be reasons why being an instance of creating artificial life is a wrong-making feature. That is, the reasons would have to make sense of why creating artificial life is always, or at least typically, *pro tanto* wrong. If (a) is interpreted along the lines of (T) the claim would be that the whole set of actions that fall under the type 'creating artificial life' is wrong – or rather, that there are reasons for this conclusion (or against the conclusion that it is right, or obligatory).

Whether we go for the (F)- or the (T)-interpretation, the burden to be lifted for someone who wants to argue that creating artificial life *is* morally significant is a fairly heavy one if DPS' definition is used. What need to be provided are reasons why it is intrinsically wrong to create artificial life, or whether the practice as such is wrong. I do not doubt that some critics hold such a view. But they certainly need not hold it. In particular, critics may hold the much more limited view that creating artificial life is *conditionally acceptable*. Defending conditional acceptability means arguing that creating artificial life is acceptable under some conditions, but not under others.

Some will no doubt object that conditional acceptability is equivalent to the claim that creating artificial life is *not* morally significant, since not all instances of it are problematic. I think this objection is misguided, for three reasons:

(1) The objection is plausible if significance is interpreted along the lines of (F). For many, it is hard to see how being an instance of creating artificial life could be a morally significant feature if only some instances of the action are (even *pro tanto*) objectionable. For in that case, surely it is not *in virtue of* being an instance of creating artificial life that the action is objectionable, but in virtue of the other features that make up the conditions under which it is unacceptable. But if significance is instead interpreted as (T), the moral judgment pertains to the whole, and the claim that each instance is wrong is not needed.

(2) The notion of conditional acceptability seems to be indispensable for thinking about technology in general. Technologies are tools that allow us to do certain things that we would not otherwise be able to do, and given that technologies are developed purposefully by human beings, we should expect them not to be unconditionally bad. Ignoring conditional acceptability in technology ethics induces thinking in terms of a dichotomy between pro- and anti-technology positions. Since pure anti-technology positions are rarely

plausible, this amounts to stacking the deck in favour of a pure pro-position that is rarely informative.

(3) Conditional acceptability entails, or provides the basis for, moral judgments in cases where the agent has to form an opinion on the practice as a whole. It entails that the practice ought not to be unconditionally permitted, and, for some conditions, that certain regulations or institutional embeddings of the practice should be put in place. And it provides the grounds for ultimate rejection of the practice for some agents. In particular, agents that have no control over whether the conditions will be met may rationally reject the practice, especially if they are not optimistic about the likelihood that conditions will be met. For such agents, the decision whether to support the practice is effectively a decision under uncertainty. For such decisions, the fact that a bad outcome – in this case instances of the practice that do not meet the acceptability conditions – *need not* occur is beside the point.

A different objection to conditional acceptability is that it threatens to make *every* action type morally significant, since (plausibly) every action type has a token that is wrong. So the reasons for judging the practice conditionally acceptable only should be suitably specific to the practice. In our case, it is particularly important those reasons do not merely result from the fact that creating life is a species of the genus *dealing with living things* – i.e. that the reasons for conditional acceptability are not at work whenever we are dealing with organisms.<sup>2</sup>

I conclude, then, that moral significance should include cases of conditional acceptability. In the next section, I will argue that the arguments for moral significance that DPS consider fare better when interpreted as arguments for conditional acceptability, and that the reasons for conditional acceptability are suitably specific.

---

<sup>2</sup> Thanks to an anonymous reviewer for *Bioethics* for suggesting that the general category of dealing with life might be the important contrast in this case.

#### **4. Arguments for moral significance**

So far I have argued that DPS' interpretations of artificial life and of moral significance are both too narrow. A thorough examination of the moral significance of creating artificial life should consider reasons for significance that stem from the design aspect of artificiality, and it should consider the possibility that creating artificial life is conditionally acceptable. In this section, I will (unfortunately all too briefly) try to show how some of the main arguments fare in light of my widened understanding of what moral significance for artificial life entails. I divide the arguments into two groups – one concerning the value of life (and nature), the other concerning uncertainty.

##### **4.1 The value of life**

Most of DPS paper discusses various arguments that aim to show that creating artificial life will, may or does already undermine the value (or moral status) of life and/or nature. Very roughly, there are two worries; (i) that creating artificial life may lead to a situation wherein some living beings' proper moral status is undermined; (ii) that creating artificial life expresses objectionable attitudes to (or valuations of) life and nature.

Consider first (i). The main worries are (a) that our creating life may produce a conception of life that undermines the moral status of living things, and (b) that we are likely to disregard the interests of living beings that we have created simply because we are their creators. On a charitable reading the arguments are presenting various scenarios that (they believe) are among the possible consequences of creating artificial life, and that are to be avoided. The most plausible direct conclusion would be to demand that certain safeguards be put in place to avoid these scenarios, such as ethics education and continued ethical evaluation of new developments (both suggested by the PCSBI (2010, pp. 134-140)). In other

words, a charitable reading would understand these as arguments for conditional acceptability. And the reasons for this are specific to creating life – it is the fact that we are creating life that (according to the objection) may lead to our maltreatment of living things in ways (a) or (b). It is of course true that any instance of dealing with living beings requires that those beings be treated right. But that is not the claim made here; instead, the claim is that the practice of creating life may result in our treating living things badly. And the condition attached to acceptability is not that living beings in fact be treated right, but that we institute mechanisms that ensure that creating life will not have those results.

It also seems that the best version of (i) would take issue primarily with the design aspect of artificiality. The fact that we can design organisms to suit our every desire seems to be a more plausible cause of the undermining of moral status of living beings than the fact that we can fabricate copies of existing life forms. And with respect to the artificial organisms themselves, a plausible worry is that their interests will not be considered in the design process; i.e., we will choose traits that suit ourselves, but not the created organisms (as is arguably already the case for many agricultural animals). This further reveals another facet of conditional acceptability – namely that creating artificial life is acceptable only if we adequately consider the interests of created organisms when doing so.

Now consider (ii), the complaint that creating artificial life expresses objectionable attitudes towards life or nature. DPS' answer to this objection is that (a) what attitudes are expressed depends on the internal mental states of the creator, and (b) not every instance of creation of artificial life expresses objectionable attitudes.

I think (a) misses the point of the objection. An expressivist critic can argue that the attitudes expressed do not depend on internal mental states. Rather, actions express atti-



tudes by revealing them – attitudes are what *makes sense of* the actions.<sup>3</sup> In our case, the idea is typically that creating artificial life reveals a desire to develop “an expansive, even limitless, ability to shape life and the future” (Baertschi, 2012, p. 438), which is inconsistent with showing “respect towards what has evolved, instead of it being objectified, instrumentalized, commodified, subjected and manipulated” (Swierstra & Rip, 2007, p. 16). Notice that it seems to be design more than fabrication that provides the basis for these concerns. Furthermore, the concerns are about *too much* design, which explains why critics do not object to domestication, but do object to synthetic biology (or, more often, to genetic modification in general).

Even though I think (a) is false, (b) is plausibly true. Not every instance of creating artificial life will plausibly express problematic attitudes, even on the objection’s own terms. Proponents of the expressivist objection typically worry about one of two things: (1) that new technologies are *replacements for* natural phenomena (Lee 1999; Preston 2008), or (2) that new technologies attempt to solve problems (e.g. environmental problems) by utilizing the same control-ideology that produced those problems (Sandler, 2007, Chapter 6). In either case, some instances of creating life plausibly are not problematic. In the case of (1), it seems that only ‘nature replacing’ is problematic. Many instances of creating life within synthetic biology would leave nature untouched. A plausible version of (1) would have no quarrel with such instances.<sup>4</sup>

---

<sup>3</sup> This is a very basic sketch of an expressivist view. For excellent statements of expressivism (in this sense), see Anderson (1993) and Anderson & Pildes (2000). Note that Anderson and Pildes do not defend an expressivist critique of synthetic biology (or other technologies).

<sup>4</sup> It should perhaps be noted that the specifics of why modern biotechnologies are nature replacing leave it somewhat murky whether *every* instance of creating life would be nature replacing. See Lee (1999) and Preston (2013).

For (2), the obvious conclusion seems to be that creating life is conditionally acceptable; genetic modification technology is acceptable so long as it is supplemented by other strategies for solving the relevant problem (such as efforts to reduce human strain on the environment), and does not repeat old mistakes of overconfidence and recklessness. In this case, the reasons for conditional acceptability are not specific to creating life, but applies to all 'technological fixes'. However, the reasons do not apply to dealing with life in general.

## **4.2 Uncertainty**

The uncertainty objection, as DPS describe it, says that creating artificial life "involves overstepping the limits of human knowledge, thus unwarrantedly risking unintended negative consequences" (Douglas et al., 2013, p. 390). The basic worry, then, is that our knowledge of biological systems and of the interactions between organisms and their environment (be that an ecosystem or a human body) is too limited for us to successfully predict what the effects of a newly created organism will be upon that environment.

DPS have a very general response to the objection, namely that the risk posed by an organism depends on its non-genealogical, especially causal, properties, and not on the way it was produced. While this is true, it does not address the connection between creating – in particular, designing – new life forms and uncertainty. The objection is about the limit of our knowledge of biological systems. In other words, it is about whether we can predict what non-genealogical, causal properties the life forms we create will have. The problem is supposed to be that we do not know enough (a) to be sure that each new organism-type is exactly the way we intend it to be, and/or (b) to be sure that it will not have harmful effects on the environment in which it will be placed (even if it is as we intend it to be).

The soundness of the uncertainty worry will depend on views about what the proper level and kind of evidence for safety is. But supposing that it is sound, the most plausible

conclusion to draw from it is not wholesale abandonment of synthetic biology. Rather, the objection supports regulation and caution, especially with the aim of avoiding interaction between artificial organisms and humans or the environment until we have sufficient evidence that they are harmless. In other words, conditional acceptability is supported. And the reasons for conditional acceptability are closely connected to the fact that we are creating life. If knowledge of type (a) is what is lacking, only instances of designing life would be problematic. If type (b) is at issue, then all instances of releasing designed organisms will be problematic, but also some instances of placing non-designed life in new environments. Full specificity is only achieved in case (a), but even in case (b) conditional acceptability does not obtain for dealing with life in general.

## **5. Conclusion**

I hope to have shown above that attention to the different conceptions of artificial life is important for understanding the content of the various objections that have been raised to creating artificial life, and that the aspect of design grounds prominent objections. Furthermore, I have argued that the strongest versions of these prominent objections argue for the conditional acceptability of creating artificial life. I have not here argued in any detail for or against either the value-of-life or the uncertainty objections, but merely shown that one way of dismissing them is unsuccessful. I hope, however, to have shown what the discussion of these objections ought to be about if it is to be fruitful, namely the question of when – under what conditions – it is acceptable to produce artificial life forms.

## **References**

- Anderson, E. (1993). *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.
- Anderson, E. & Pildes, R. (2000). Expressive Theories of Law: A General Restatement. *University of Pennsylvania Law Review*, 148(5), 1503-1575.
- Baertschi, B. (2012). Defeating the Argument from Hubris. *Bioethics*, 27(8), 435-441.
- Bedau, M., Church, G., Rasmussen, S., Caplan, A., Benner, S., Fussenegger, M., Collins, J., & Deamer, D. (2010). Life after the Synthetic Cell. *Nature*, 465, 422-424.
- Boldt, J., & Müller, O. (2008). Newtons of the Leaves of Grass. *Nature Biotechnology* 26, 387-389.
- Douglas, T., Powell, R., & Savulescu, J. (2013). Is the Creation of Artificial Life Morally Significant? *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 688-696.
- Endy, D. (2005). Foundations for Engineering Biology. *Nature*, 438, 449-453.
- Goodpaster, K. (1978). On Being Morally Considerable. *The Journal of Philosophy*, 75(6), 308-325

- Holm, S. (2012). The Scientific Aspirations of Synthetic Biology and the Need for Analytic Ethics. *Ethics, Policy & Environment*, 15(1), 25-28.
- Lee, K. (1999): *The Natural and the Artefactual: The Implications of Deep Science and Deep Technology for Environmental Philosophy*. Lanham, MD: Lexington Books.
- Machery, E. (2012) Why I Stopped Worrying about the Definition of Life... and Why You Should as Well. *Synthese*, 185, 145-164.
- O'Malley, M., Powell, A., Davies, J.F., & Calvert, J. (2008) Knowledge-Making Distinctions in Synthetic Biology. *BioEssays*, 30(1), 57-65.
- Presidential Commission for the Study of Bioethical Issues (PCSBI) (2010). *New Directions: The Ethics of Synthetic Biology and Emerging Technologies*. Washington, DC: PCSBI.
- Preston, B. (2013). Synthetic Biology as Red Herring. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 649-659.
- Preston, C.J. (2008). Synthetic Biology: Drawing a Line in Darwin's Sand. *Environmental Values*, 17(1), 23-39
- Sandler, R.L. (2007). *Character and Environment: A Virtue-Oriented Approach to Environmental Ethics*. New York, NY: Columbia University Press.

Sperber, D. (2007). Seedless Grapes: Nature and Culture. In E. Margolis & S. Laurence (Eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford: Oxford University Press.

Swierstra, T., & Rip, A. (2007). Nano-Ethics as NEST-Ethics: Patterns of Moral Argumentation About New and Emerging Science and Technology. *NanoEthics*, 1(1), 3-20.

## Article 2 Similarity Arguments in the Genetic Modification Debate

### 1. Introduction

In debates on genetic modification (GM)<sup>1</sup>, references to the similarity between these new technologies and existing, well-established technologies are common. For example, an editorial in *Scientific American* arguing against labelling of GM food makes the following claim:

We have been tinkering with our food's DNA since the dawn of agriculture. By selectively breeding plants and animals with the most desirable traits, our predecessors transformed organisms' genomes, turning a scraggly grass into plump-kerneled corn, for example (Labels for GMO Foods, 2013)

I will refer to this basic claim – that GM is in fact similar (in some relevant respect) to an established technology – as a *similarity claim*. Often, as in this case, the specific import of the similarity claim is not made explicit, although it is clear from the context that the fact is supposed to play some part in defending GM from critics. I will refer to the use of similarity claims to defend GM as a *similarity argument*. Similarity arguments are ubiquitous. Tsjalling Swierstra and Arie Rip (2007, p. 9) identify it as a common “trope” of debates about new and emerging technologies. In a study using facilitated discussions between scientists and non-scientists on GE (and in particular the common objection that GE is unnatural), at least

---

<sup>1</sup> By ‘genetic modification’ I mean technologies that directly alter the genomes of organisms, such as transgenesis, genome editing and synthetic biology. Several ‘borderline’ technologies could be included, e.g. mutagenesis and protoplast fusion. These are rarely discussed, and I will not say anything specific about them in the paper. I am mainly concerned with the debate concerning GM applied to plants and microorganisms, but I presume similarity claims and arguments occur in the case of human- and animal GM as well – and indeed in many other fields of applied ethics.

three of the five scientists used the similarity claim as their basic argument (Deckers, 2005). And Jonathan Pugh (2015a) describes it as “the standard response” to the so-called playing God objection.<sup>2</sup>

Despite its ubiquity, it is less than clear what the precise *content* of the argument – its premises and conclusion – is supposed to be. Pugh gives this concise statement of the argument: “Humans have been selectively breeding both plants and animals for hundreds of years, and this can be viewed as an indirect form of genetic modification that we do not find morally problematic” (Pugh 2015a). It seems that the fact that we do not find selective breeding morally problematic, even though it, too, is a form of genetic modification, is an important premise. Nils Holtug suggests the same when he argues that any sense of ‘natural’ that is to function as an objection to GM based on its being *unnatural* “should not apply to the intentional genetic modification of organisms through selective breeding (since even proponents of the ... objection want to allow this technology)” (Holtug, 2009, p. 237). Swierstra and Rip suggest that the similarity argument is a kind of argument from precedent: “If we see these earlier technologies as being in accordance with our present moral intuitions, we should now be consistent and see the new technologies as similarly acceptable” (Swierstra & Rip, 2007, p. 9). R. Paul Thompson similarly argues that “[v]ilifying this or that domain of science and technology (GM agriculture, for example) while accepting the benefits in another domain (GM medicine, for example) is ... inconsistent” (Thompson, 2011, p. 110).

Similarity arguments, then, are simultaneously ubiquitous, *prima facie* plausible and somewhat unclear. My aim in this paper is to provide some clarity. My strategy is to (re)construct an argument, containing premises and a conclusion, based on the claims made in the quotes in the preceding sections. I suggest that we can distinguish between a weak and a strong version of this argument, depending on what the conclusion is supposed to be.

---

<sup>2</sup> In a different version of the same article (Pugh, 2015b), it is described merely as “familiar”.



I then evaluate the argument's persuasiveness. My aim is not to refute or even criticize the argument, but rather to show which parts of the argumentative space the argument closes off to the GM critic and which it does not – that is, what kinds of objection to GM similarity arguments are effective against and what kinds they are not effective against. I then apply this analysis to two specific instances on the similarity argument found in the GM literature.

## **2. Similarity Arguments**

When talking about similarity arguments, we can make a distinction between *specific* instances and a *generic* argument (or argument structure). A specific instance of the similarity argument includes reference to a particular *objection* to GM and to a particular existing *technology*. In the previous section, we have mainly seen similarity arguments directed at so-called intrinsic objections, i.e. objections according to which there is something problematic in the very fact of modifying organisms directly. These include complaints that GM is unnatural, that it constitutes our 'playing God', that it commodifies or otherwise degrades living beings, or that it is an expression of an objectionable attitude of domination over nature – as well as brute complaints that genetic modification is simply wrong. Less frequently, similarity arguments are used against other objections, such as that GM creates risk to human health and the environment, or even that GM leads to a concentration of market power. In this paper, however, I focus on the intrinsic objections since these are by far the most common targets of similarity arguments.

Besides a particular objection, a specific similarity argument also invokes some particular technology with which GM is compared. The most frequently used technologies are technologies whereby we also alter organisms on the genomic level, i.e. everything from early domestication through scientific breeding and hybridization. Another popular comparison is medicine, both taken as the general practice of healing the sick and combatting

disease, and GM technologies in medical applications (e.g. the production of insulin in modified bacteria). What technology is used obviously depends on the objection to GM that the similarity argument targets. The comparison technology must share the feature that grounds the objection. Thus medicine as a general practice is typically invoked to combat objections based on GM being unnatural, while scientific breeding and hybridization are used to counter objections based on objectionable attitudes to life or nature.

In this paper I will focus on two similarity arguments. The first argument is directed at the unnaturalness objection as it is commonly expressed by members of the general public. Similarity arguments using both medicine and non-GM agriculture has been used against this objection. The second is a different variety of the unnaturalness objection, which is mostly found among philosophical critics of GM. This strand of objection is typically based on the claim that GM turns organisms into or treats organisms as artefacts. The relevant comparison technologies are those by which humans have previously changed the genomic make-up of organisms, such as domestication/cultivation and scientific breeding.

In the remainder of this section and in the next section, I will focus on the generic similarity argument. The generic similarity argument is the general argument structure that specific similarity arguments exemplify. It has two open variables, namely the two things just discussed: the established technology that GM is similar to (call it T) and the feature, used by the GM critic as a basis for her objection, with respect to which the two technologies are similar (call it F). The generic argument consists in two premises and a conclusion. The first premise is the basic similarity claim

(1) F is a feature of T as well as of GM

The second premise is that the T in question is not morally problematic. There are several different ways of framing this. Pugh argues that selective breeding is “an indirect form of genetic modification that we do not find morally problematic”; Holtug stresses that even

those who object to GM “want to allow” selective breeding; Swierstra & Rip describes the claim as being that T is “in accordance with our present moral intuitions”; and Thompson focuses on the fact that critics “accept the benefits” of GM in medicine. I will abstract from these differences in framing and state the premise as

(2) T is not morally problematic

Finally, the conclusion of the argument is that the objection to GM based on F is in some way in trouble due to (1) and (2). In many cases, especially in public debate, the precise content of the conclusion is not made explicit – (1) and (2) are simply put forward as a rejoinder to the F-based objection. The examples cited from the philosophical literature give differing conclusions as well. Swierstra & Rip interpret similarity arguments as drawing the conclusion that “we should ... now see the new technologies as similarly acceptable” (Swierstra & Rip, 2007, p. 9); Pugh introduces his similarity argument to exemplify why the playing God objection is “not persuasive” (Pugh 2005a); Thompson argues that the argument shows the critic to be “inconsistent” (Thompson, 2011, p. 110); and Holtug claims that it is one among several “requirements” of any naturalness based objection that the sense of ‘natural’ used does not apply to selective breeding (Holtug, 2009, p. 237). Roughly speaking, we can distinguish between a weak and a strong conclusion. According to the weak conclusion, the GM critic owes an explanation for why F is a reason to find GM morally problematic when it is not such a reason in the case of T. According to the strong conclusion, the critic’s objection is simply refuted by the similarity argument. That is, the argument has shown that F is not a reason to find GM morally problematic. We can state these two conclusions as follows:

The weak conclusion (WC): So the GM critic owes an explanation of why F is a reason to find GM morally problematic, but not a reason to find T morally problematic

The strong conclusion (SC): So F is not a reason to find GM morally problematic

It is worth noting that the generic similarity argument is my reconstruction of a pattern of argument that is ubiquitous in the debate on GM, but whose precise content and import is often, if not always, left unclear. The procedure I am using is thus somewhat different from what many might think is ‘best practice’ in philosophy, namely finding the best and most well-developed instance of a view or an argument and criticising that. I diverge from that practice because for two reasons. First, the similarity arguments that are ubiquitous in the GM literature are of the vague and somewhat hasty style I have described above, and very few (if any) thoroughly worked-out versions exist. The combination of ubiquity and vagueness suggests to me that clarification is called for. Second, my aim is not to argue that similarity arguments are bad arguments, but to investigate *how* effective they are against common objections to GM. Consequently, I do not believe that I am criticisable for arguing against a straw man, or of criticising an argument that I have simply made up.<sup>3</sup>

### **3. Evaluating the generic similarity argument**

I have argued that we can make a distinction between a weak and a strong version of the generic similarity argument. Either version is a reply to an objection to GM of the form “F is a reason to find GM morally problematic”. The weak argument is as follows:

(1) F is a feature of T as well as of GM

(2) T is not morally problematic

(WC): So the GM critic owes an explanation of why F is a reason to find GM morally problematic, but not a reason to find T morally problematic

---

<sup>3</sup> Thanks to Julian Savulescu and an anonymous reviewer for *Ethical Theory and Moral Practice* for pressing me to clarify this.

The strong argument substitutes (SC) for (WC), yielding:

(1) F is a feature of T as well as of GM

(2) T is not morally problematic

(SC): So F is not a reason to find GM morally problematic

In this section, I critically evaluate these generic arguments. I argue that there are two ‘gaps’ between the two premises and the conclusion(s), and suggest how these gaps may be filled. I will further argue that there is a clear risk of fallacies of equivocation with respect to what is meant by ‘morally problematic’ and what the feature, F, that grounds the objection to GM is.

### **3.1 The first gap: From all-things-considered judgments to *pro tanto* reasons**

The first gap concerns a discrepancy between how (2) is formulated and how (WC) and (SC) are formulated. According to (2), T is not morally problematic, while (WC) and (SC) both make reference to the claim that *F is not a reason* to object to GM. In order for the argument to work, we need the claim

(3) F is not a reason to object to T

There are three possible ways of getting to (3). First, we can attempt to *infer* (3) from (2). Second, we can argue that the relevant critic of GM accepts (3) rather than merely (2). And third, we can reframe the similarity argument and substitute (3) for (2) as a premise.

Consider the first option. Inferring (3) from (2) would require defending the claim

(4) If T is not morally problematic, then F is not a reason to find T morally problematic.

As a general principle, (4) fails since it does not acknowledge the distinction between *pro tanto* reasons and all-things-considered judgments. This distinction is a familiar one. A *pro tanto* reason is a consideration that counts in favour of (or against) a particular conclusion.

But there may be other considerations that provide reasons for the opposite conclusion. All-things-considered judgments are the result of weighing up all the *pro tanto* reasons that pull in different directions. There is no guarantee that the judgment that a given *pro tanto* reason favours is also the judgment that is the all-things-considered right one. For example, even if using GM entails some risk of harm to health or the environment, the benefits of the technology may be sufficient to outweigh this bad aspect. Consequently, we cannot infer from the fact that the judgment recommended by a *pro tanto* reason is not the all-things-considered right one that the *pro tanto* reason does not hold. So (4) is not true as a general rule. But (3) could still be inferred from (2) in a specific case if either of two conditions are met: (a) F is not a *pro tanto* reason only, but rather a *sufficient* reason; (b) T does not have any other feature (or set of features) that provides a reason for judging that T is *not* morally problematic (or any such reason is or would be weaker than F, and would thus be outweighed by F).

Consider first (a), that F is a sufficient reason rather than merely a *pro tanto* reason. A sufficient reason is (as the name suggests) a reason that is *in itself sufficient* for securing the judgment. Sufficient reasons are typically associated with fairly strong forms of deontology that issue exceptionless prohibitions against performing actions with certain properties (e.g. actions that are lies, or the killing of a person, or a violation of rights). For objections to GM that rely on such absolutist views – i.e. views according to which anything that has the feature F is all-things-considered morally problematic merely in virtue of having F – (3) follows from (2), and thus the first gap in the generic similarity argument is filled. Although I suspect at least some critics do rely on absolutist views, not everyone does. In many cases, it will be a matter of interpretation whether the GM critic in question takes F to provide a *pro tanto* or a sufficient reason for finding GM morally problematic. In a subset of these, there is no concrete critic at all, since the similarity argument is used to counter abstracted versions

of common objections. I will take up the question of how we should interpret objections to GM in the final section of the paper

Now consider (b). Even if F is only a *pro tanto* reason, the judgment that T is morally problematic would still follow if there were no stronger reason on the other side. For many of the established technologies used in common similarity arguments, such as medicine or domestication and cultivation – there clearly are strong reasons against the judgment that they are all-things-considered morally problematic, such as the large benefits to human welfare these practices entail or have made possible. The best case for making an inference of type (b) would be to invoke a comparison technology or practice that offers no benefits, but still has the feature F. It is, of course, a common strategy in applied ethics to imagine cases where a feature of interest is ‘isolated’ in this way and elicit intuitions about it. But similarity arguments are not arguments that use hypothetical cases, but arguments that use actual and well-established technologies – and much of their *prima facie* persuasiveness relies on this fact. I doubt, for that reason, that the (b)-strategy for inferring (3) from (2) would be successful for any actual similarity argument. But of course, if a concrete critic of GM holds that the relevant T possesses no or too few features that provide reasons in favour of it, (3) would be established.

I now return to the two other ways of getting (3), besides inferring it from (2). The first of these is to argue that the relevant GM critic actually holds (3) rather than merely (2) – that is, that she already believes that F is not a reason to find GM morally problematic. Whether this is the case will depend on the concrete critic in question, and again the issue of interpretation arises. But the fact is that if the critic actually holds (3), the first gap disappears.

The final possibility is to reformulate the similarity argument by substituting (3) for (2), and hence to make the claim that F is not a reason to find T morally problematic a prem-

ise in the argument. This does not seem to me to be the best interpretation of the examples of similarity arguments from philosophers given above, but it may be a plausible interpretation of many uses of similarity arguments in public debate. There is certainly no in-principle reason why similarity arguments could not be framed in this way. If they are, they are strengthened in the sense that the first gap never opens up. However, this comes at a price, since the intuition that a certain well-established technology is not all-things-considered morally objectionable is typically very strong, while the intuition that a certain feature of such a technology does not provide a *pro tanto* reason for finding it morally problematic is less strong. Still, it seems plausible that many have the intuition that, for example, the fact that domesticated plants and animals are genetically altered relative to their wild ancestors does not provide even a *pro tanto* reason to object morally to the practice of domestication.

### 3.2 Second gap: Transferability of reasons

Suppose that we accept that F is not a reason to find T morally problematic, i.e. that (3) has been established. There is now a second gap between (3) and the conclusions

(WC): So the GM critic owes an explanation of why F is a reason to find GM morally problematic, but not a reason to find T morally problematic

and

(SC): So F is not a reason to find GM morally problematic

In both conclusions, the assumption is that there is some demand that, if you judge F to be a reason in one case, you ought also judge it to be a reason in the other case. Let me first consider the strong conclusion (SC). Here, we need some way of getting from the fact that F is not a reason to find T morally problematic to the conclusion that F is not a reason to find GM morally problematic, i.e. from (3) to (SC). Unlike what was the case with respect to the first gap, we cannot even hope to find textual evidence that the critic in question accepts



(SC), except in cases where the critic blatantly contradicts herself. Neither can we rely on a direct intuition that (SC) is true, at least without making the similarity argument redundant. So we are left with the option of inferring (SC) from (3). Doing so would require defending the claim

(5) If F is not a reason to find T morally problematic, then F is not a reason to find GM morally problematic

In other words, we must be able to *transfer* the reason-giving force of F from the case of T to the case of GM

Transferability as a general principle has come under attack by those attracted to particularism in ethics. A prominent example is Jonathan Dancy, who criticizes what he calls “switching arguments”:

A switching argument is an attempt to determine what to say here by appeal to what we say about something else ... The particularist supposes that all these argument forms are attempts to *force* us into a view of the present case by appeal to some feature of another case, or some comparison between this case and that. And the response to moves of that sort is invariably sceptical; they all attempt to pre-empt the authority of the present case (Dancy, 1993, p. 64-65).

It should be noted here that Dancy allows that “our judgment can be enlightened by a comparison between a new case and others in our experience”. What he objects to is the idea that there is a rational *demand* to conform our judgments about one case to our judgment in another.

Counting in favour of transferability is the fact that the use of data from one case to draw conclusions about another case is very widespread in practical reasoning. Conse-

quently, rejecting its validity is would place us in a situation of near paralysis with respect to our ability to make decisions based on reasons and reasoning. Thus Simon Blackburn:

[A] great deal of ethical thought ... concerns how to act. But such thought is essentially a matter of selecting and weighing features that recur from one situation to another. In trying to discover what to do, we imagine different actions, and register their good and bad features. It is essential to this process that that these features are reliably extracted from any contexts or total situations in which we have come across them, and carry some moral import when transplanted into the new hypothetical situation ... If these features lost their moral import just as soon as they were abstracted from other cases, in which they had been marinated with others to give some holistic moral gestalt, this process would be totally unjustified (Blackburn, 1996, p. 97)

I think Blackburn is right in pointing out the potentially far-reaching implications of denying transferability. Furthermore, as Blackburn suggests, denying transferability conflicts with the phenomenology of ethical thinking. We do in fact take the things that matter to matter generally. And we do come to judgments about new cases using reasons judgments that we accept antecedently, rather than simply look at the case and directly intuit a judgment. Any limitations to the transferability of reasons should leave our ability to reason about cases using beliefs about reasons with general scope intact.

Interestingly, there is a weaker version of the worry expressed by Dancy that does exactly that. On this view, reasons should be understood as *defaults* (Horty, 2007; Horty, 2012). A default is here understood as a consideration that supports a judgment, but which

may be outweighed (as above) or *excluded* altogether by other considerations.<sup>4</sup> It is this latter option that creates a problem for transferability, since a default that is not excluded in one case may nevertheless be excluded in another case. The difference between being outweighed and being excluded is that an outweighed reason still retains its reason-giving force, whereas an excluded reason does not. For example, the fact that an action of mine puts you in some danger of harm – say, I try to knock a can of soda out of your hand with a baseball – is a reason not to perform that action. But the fact that you consented to my performing the action – say, you dared me to try to knock the can out of your hand – removes the reason-giving force of hurting you altogether. And this is different from the case where there were merely strong counterweighing reasons for trying to knock the can out of your hand – say, that it was attracting wasps that threatened to sting you. Here, the possibility of hurting you with the baseball still retains some reason-giving force, as witnessed by the fact that I would be obliged to apologise if you were in fact to get hurt.

So if a feature is a reason for some judgment in one case, it may nevertheless not be a reason for that same judgment in another case – if it is excluded in the latter case. This means that we can alleviate Blackburn’s worry that reason-giving features lose their reason-giving force as soon as they are abstracted from the case in which we originally found them to have moral import, but keep the possibility open that the reason-giving force does not transfer to this *particular* new case. In the standard case, we have established that some consideration is a default reason and then attempt to show (or disprove) that the default reason holds in some particular case. In the case of the similarity argument, the situation is slightly

---

<sup>4</sup> Or more precisely, other considerations may provide a reason to judge that the original consideration is not, after all, a reason for the original judgment. The notion of exclusion (and enabling) that is the basis of the default view also constitutes Dancy’s grounds for defending particularism. It is thus plausible to take the default view to capture all the worries about transferability that Dancy’s view does.

different. The claim in (5) is that F is *not* a reason in the case of T, and that this *lack* of reason-giving force transfers to the GM case. So in order to be able to infer (SC) from (3), the proponent of the similarity argument would have to make the case that the lack of reason-giving force of F in the case of T is not due to its merely being excluded in that case. Furthermore, the SA must make the case that there are no special *enabling* considerations in the case of GM that endows F with reason giving force in this case although it has no such force normally.<sup>5</sup> Unless the proponent argues for these, (SC) has not been shown to be true – there remains space for the GM critic to hold a view whereby F is a reason to find GM morally problematic, but not a reason to find T morally problematic.

This brings me to the weak conclusion (WC), that the critic owes an explanation of why F is a reason to find GM morally problematic, but not a reason to find T morally problematic. The discussion above shows how the critic may come up with such an explanation. The difference can be justified either by showing that there are enabling considerations that make F a reason in the GM case, or that there are excluding considerations that remove the reason-giving force in the case of T. A lot now hinges on what we mean by the claim that the critic “owes an explanation”. There is a spectrum of progressively stronger claims here. The mildest claim would merely be that it would improve the critic’s view if she were to present an explanation of the seeming inconsistency. The strongest claim would be that the similarity argument has established a presumption in favour of the *strong* conclusion that F is not a reason to find GM morally problematic. If the mild form of (WC) is what is meant, it is fairly uncontroversial that (WC) follows from (3). The strongest version, that there is a presumption in favour of (SC), on the other hand, seems to me to require at least that the

---

<sup>5</sup> The relation between excluding considerations and enabling considerations is unclear – they may simply be different ways of describing the same thing. My sense is that the main difference between considerations that are most aptly described as excluders and enablers, respectively, is whether the situations they point to can be described as ‘standard conditions’ or the ‘normal’ state of affairs.

proponent of the similarity argument considers some possible contextual differences between the cases – i.e. considers what possible exclusionary or enabling considerations may exist. But generally, the question of when a presumption in favour of a certain claim exists or has been established is underexplored in ethics, and I offer no theory of it here.

### 3.3 Equivocations

The third and final potential issue with the similarity argument concerns possible equivocations in premises (1) and (2). In (1) the F that is a feature T may not be the same as the F that GM critics take to be a reason finding GM morally problematic. In (2), the sense in which T is not morally problematic may not be the sense in which GM is supposed to be morally problematic.

There are three particular issues that the proponent of the similarity argument should be aware of. First of all, the meaning of ‘morally problematic’ is not as obvious as one might be tempted to think. Commonly, no precise meaning of ‘morally problematic’ is stated by the critic, especially where the critic is a layperson (or, of course, where the critic is merely imagined). The sense of ‘morally problematic’ that is typically understood by proponents of the similarity argument is ‘morally unacceptable’. But there are at least two possible weaker senses. First, ‘morally problematic’ may mean ‘*pro tanto* morally problematic’ (similar to W.D. Ross’ notion of a *prima facie* duty). This sense is very closely related to what I discussed above under the heading of *pro tanto* reasons. There is only little difference (if any) between saying that F is a *pro tanto* reason to judge that GM is morally problematic (all-things-considered), and saying that F is a sufficient reason to judge that GM is *pro tanto* morally problematic. So in effect, this possibility reduces to the possibility that F is merely a *pro tanto* reason. The second weaker sense of ‘morally problematic’ is ‘not unconditionally morally acceptable’ – or, more colloquially (though less precisely) ‘conditionally acceptable’.

The conditions referred to could be various regulatory mechanisms, such as safety guidelines or oversight measures; they could be social or economic background conditions; or they could simply be the existence of sufficiently strong reasons to engage in the relevant application of GM. There is some evidence that public attitudes to GM are well captured by the notion of conditional acceptability (Marris et al., 2001).

The second issue concerns F. Setting aside the possibility of gross misinterpretations of objections, equivocations on F can occur if the specific sense of F that the GM critic suggests is a reason to find GM morally problematic (call it  $F_{GM}$ ) is a species of the more general kind F. So while it is true that F, in *some* sense, holds for T,  $F_{GM}$  does not. For example, the feature ‘that X constitutes the rearranging of naturally occurring materials’ is very general. It is probable that the GM critic has something more specific in mind – say, rearranging naturally occurring materials in a certain way, or rearranging a certain kind of naturally occurring materials. Once more, careful interpretation of critics’ claims is needed.

The third issue is related to the second. Even where F does refer to the same phenomenon, it may be the *amount* or *degree* of F that does the reason-giving work. That is, the critic might not believe that the mere fact that GM has the feature F is problematic, but that it is the fact that it has *too much* of F that is. There are clear cases where degree-differences make all the difference to the reason-giving force of a feature. For example, monetary cost’s reason-giving force is clearly sensitive to degree difference. As is causing harm in at least some cases, e.g. in the case of proportionality judgments in the use of military power. These are cases of trade-offs, where some amount of a bad thing (e.g. monetary loss, deaths) is accepted in order to obtain some good (e.g. a nice dinner, lasting peace in the Middle East). Critics of GM may have such trade-offs in mind. The argument would then be that *given* the (perceived) benefits of (an instance of) GM, the degree of F involved is reason to judge that we ought not engage in (this instance of) GM. In other words, *this* degree of F is too high a price

to pay given the benefits of GM. Note the similarity to the issue of *pro tanto* vs. all-things-considered reasons. In that case, the issue was whether T had features that (could) outweigh the weight of the reason provided by F. Here, the issue is whether a different level of F requires stronger counterweighing benefits – and, supposing that the level of F is indeed higher for GM than for T, that the reasons in favour of GM need to be weightier than those in favour of T.

But it may not be the case that critics have trade-offs in mind. Rather, they may think that *some* degree of F is unproblematic, but that a sufficiently high degree of F is a reason to find an activity morally problematic. Some might doubt that degrees could matter for the very reason-giving force of a consideration, as it would do in this case. Indeed, it is common to see arguments to the effect that a difference between GM and an earlier technology is *merely* one of degree rather than kind, and that *therefore* it cannot make an ethical difference. A reason for such doubt might be this: Suppose that  $F_1$  (F to degree 1) is not a reason for some judgment, and that  $F_2$  is a degree of F that is only a tiny, imperceptible amount larger than  $F_1$ . By the principle of treating like cases alike, we should conclude that therefore  $F_2$  is not a reason. By repeating this reasoning, we are eventually driven to the conclusion that not even  $F_n$ , the largest possible degree of F, could provide a reason. So if there is a degree  $F_1$  that is not a reason for the judgment, then no degree of F is such a reason.

The reasoning above is similar to the reasoning that generates sorites paradoxes, such as the case of the relationship between the number of hairs on a man's head and his being bald (a man with 100,000 hairs is not bald, and *one* hair's difference doesn't make a man bald; so a man with 99,999 hairs is not bald; so a man with 99,998 hairs is not bald ... so a man with zero hairs is not bald). But note that there is no *paradox* in our case. In the case of the bald man, the paradox arises because the end point – that a man with zero hairs is not bald – is clearly false. But in our case, whether  $F_n$  has reason-giving force is contested. In-

stead, the problem with our quasi-sorites case is that denying the possibility that  $F_n$  is a reason when  $F_1$  is not entails that we are committed to a type of rational slippery slope for features that come in degrees: Once you have accepted that some degree of  $F$  is not a reason for a judgment, you are rationally committed to accepting that no degree of  $F$  could be a reason. Thus critics would seem to be *justified* in objecting to technologies using arguments that look like slippery slopes, which is a consequence that we should not be happy with.

In order to deal with this issue, the proponent of the SA can attempt to show that degrees cannot, or ought not, make a difference in the specific case at issue. There are three strategies available. First, one can simply deny that there is in fact a difference in the degree of  $F$  between GM and T. If that is the case, the difference in degree of  $F$  obviously cannot account for different judgments, since it does not exist. Second, one can argue that the degree of  $F$  in GM and T lies on the same side of the (possibly fuzzy) limit beyond which  $F$  provides a reason. An example of this phenomenon is budgetary restrictions; if either of two courses of action cost more than our available budget, it does not matter (for the reason-giving force of the budgetary restriction) that one is more costly than the other. This may require interpretation of the possible grounds a critic might have for taking  $F_n$  to be a reason. Third, one can deny that  $F$  is the kind of thing that admits of degrees, or, weaker, that any plausible reason-giving force of  $F$  cannot vary with the degree of  $F$ . There are many examples of the former, e.g. the fact that someone is dead or the fact that you signed the contract. Nozick's conception of rights is an example of the latter. According to Nozick "stealing a penny or a pin or anything from someone violates his rights", and thus no degree difference can matter to whether a right is broken or not (Nozick, 1974, p. 75).



### 3.4 Taking stock

My arguments in the three preceding sections have shown that the similarity argument leaves some argumentative space for the GM critic. First of all, it is possible for the critic to maintain that F is a *pro tanto* reason to find T morally problematic, even though T is not all-things-considered morally problematic. Second, the critic can argue that contextual differences – exclusionary and enabling considerations – make it the case that F is a reason to find GM problematic, but not to find T morally problematic. Third, the critic can argue that the sense of F that she takes to provide a reason for objecting to GM is not the same sense of F that is present in the case of T – either because  $F_{GM}$  and  $F_T$  differ in kind, or because they differ in degree. Fourth, and finally, the critic can argue that the sense in which she takes GM to be morally problematic is not the sense in which T is not morally problematic. Note that making any one of these arguments would be sufficient to block the strong conclusion of the similarity argument or to discharge the explanatory burden that the weak conclusion imposes on the critic. In the next section, I look at whether three types of objection to GM that have been subjected to a similarity argument can use this argumentative space to avoid the conclusions of the similarity argument.

### 4. Specific similarity arguments in the genetic modification debate

Recall that specific similarity arguments target a particular objection to GM using a particular comparison technology. In this section I discuss two such specific arguments. The first targets the objection based on naturalness as it exists in the public debate over GM food, and uses medicine (in general as well GM medicine specifically) as a comparison technology. The second targets a more specific version of the unnaturalness argument that finds fault with our transforming organisms into artefacts, and uses selective breeding and similar technologies as a comparison technology.

#### 4.1 Naturalness and medicine

There is evidence that a major source of public scepticism towards GM food is based on the idea that GM food is unnatural (Gaskell et al., 2010, p. 38). But medicine is also unnatural, in the sense that it consists in preventing natural process from taking their course. Similarly, specific medical treatments based on GM seem to be just as unnatural as GM food is. Insofar as we think that medicine in general and GM medicine specifically are not morally problematic, we have premises (1) and (2) of a similarity argument against the unnaturalness objection to GM food. How could a defender of the unnaturalness objection avail herself of the argumentative space left by the similarity argument?

The most promising route to take for the GM critic in this case seems to me to focus on the meaning of '(un)naturalness'. The comparison with medicine in general seems to rely on an understanding whereby everything that has been produced by human beings is unnatural. Taking unnaturalness to be a reason to object to a practice would thus make every aspect of human culture objectionable. For that reason alone it is doubtful that any critic bases her objection in this sense of 'natural'. And psychological and social scientific studies have shown that ordinary people's judgments about what things are natural (especially in the domain of food) are fairly sophisticated, and depend *inter alia* on whether things have been added or subtracted from the 'original', whether alterations are mechanical or chemical and whether the changes could in principle have occurred without human help (Rozin 2005; Rozin, Fischler & Shields-Argeles, 2009; Mielby et al., 2010). Paul B. Thompson<sup>6</sup> has suggested that the most common conception of naturalness at play in the GM food debate is an "artisanal" conception (Thompson, 2003). On this conception, natural farming is farming

---

<sup>6</sup> To avoid unnecessary confusion: Note that *Paul B.* Thompson is a different person from *R. Paul* Thompson, cited in §1.

that uses contextualized knowledge about the specific ecosystem in which it takes place and exploits the synergies it contains between plants, animals and soils. It is contrasted with 'industrial' methods such as the use of chemical fertilizer and pesticides. GM agriculture is associated with this latter, unnatural approach to farming. Interestingly, this connects the unnaturalness objection to the so-called 'technological fix'-objection, according to which the problem with GM agriculture is that it (so critics claim) seeks to solve problems caused by industrialization – such as pollution, topsoil loss and increasingly aggressive pests – with more industrialization, instead of removing the root causes of those problems (see e.g. Sandler, 2007, Chapter 6). From the point of view of the similarity argument using medicine as such, the most interesting point is that the artisanal conception of naturalness does not apply to such very general technologies (medicine as such, agriculture as such), but at most to certain ways of practicing them. It is thus not true that the general practice of healing the sick is artisanally unnatural, and hence that general practice does not share the relevant F that underlies the objection to GM.

Assuming that the artisanal conception of naturalness is the conception that underlies the objection still does not do away with the similarity argument that uses GM medicine as the comparison technology. For presumably there is an artisanal conception of natural medicine as well, and critics of GM food presumably do not take the fact that novel medicines developed by GM are non-artisanal as a reason to find them morally problematic. Can this difference in judgments be rationalized by either of the two remaining routes?

First, consider the possibility that GM medicine is *pro tanto* problematic because it is unnatural, but that this reason is outweighed by its benefits. In expressing misgivings about "certain forms of stem-cell research" G.A. Cohen argues:

And do not say, “If *your* child’s life depended on it ...” There are all kinds of awful things that I would not otherwise dream of doing that I might do if my child’s life depended on it ... When people say: “If *you* had cancer...,” one can sometimes reply: “Yes, of course, that might unbalance my judgment.” Making people imagine that they are in dire straits in order to cause them to agree with something is an attractive resort for those whose arguments are not (otherwise) strong (Cohen, 2011, p. 209).

What Cohen suggests is that people will be willing to accept more or less anything if saving lives, especially one’s child’s or one’s own, is placed in the other side of the scales. So the mere fact that we do not tend object to GM-based medicine is not a particularly strong reason to believe that there is no reasons for objecting to it (including the possible reason that it is unnatural in the relevant sense).

Second, consider the possibility that contextual considerations (excluders or enablers) can explain why ‘artisanal’ unnaturalness is a reason to find GM food problematic but not GM medicine. One plausible view is that people take natural foods to be healthier (more nutritious and/or less likely to be detrimental to health in any way) or superior in terms of taste than non-natural foods, i.e. to have properties that are desirable in foods. And the comparable belief, that natural medicine is better at curing disease, i.e. has properties desirable in medicines, is not as prevalent (and of course those who *do* think natural medicine is better are likely to deny that naturalness is no reason in the case of medicine). A study (Rozin et al., 2004) bears this out: More than 70% agreed that natural foods were more healthy than non-natural (in an artisanal sense of natural/unnatural), while only 24% agreed that natural medicines were more effective at curing diseases.

Finally, consider the possibility that the proponent of the similarity argument has misunderstood in what sense critics take GM to be morally problematic. Even if it makes sense for people to prefer artisanally natural foods to non-natural foods, this does not seem to warrant the view that GM food ought not to be produced. But there is unquestionably widespread resistance to *allowing* GM food, especially in Europe. Could it nevertheless be the case that naturalness *per se* only functions as a reason for preferring natural foods? Could the more extensive resistance be explained if that were the case? I believe it could, by three things (or a combination of two or more of them). First, unnaturalness is also tied to a belief that GM crops are risky in the sense that they are likely to harm the environment or to contaminate non-GM foods. Given such a belief, it makes sense to want to prohibit GM crops. Second, the preference for naturalness, especially insofar as it is based on the belief that natural foods are healthy, rationalizes an objection to being subjected to unnatural foods unknowingly. That is, naturalness may play a role in an argument for labelling rather than a ban. Third, support for not allowing GM crops need not be based on the claim that planting such crops is *wrong*, but could instead be understood as an integral part of the view that we should go back to a more artisanal way of farming. A critical stance towards GM is thus closely connected to a critical stance on 'industrial' agriculture, and a positive stance on such things as agroecology. It expresses a more general view about how food production should be practiced in the future.

#### **4.2 Manipulation of organisms and breeding techniques**

The second specific similarity argument targets an objection to GM that is found among philosophers rather than ordinary citizens (although some such may, of course, hold it too). According to the objection, the problematic feature of GM is that it amounts to the intentional design of organisms' genotypes, thereby making those organisms into artefacts (call

this ‘manipulation’). Manipulation may be taken to be another, rather specialized conception of naturalness. The comparison technology used is earlier ways of modifying organisms – from domestication and cultivation to scientifically informed breeding and hybridization. Again, the question is whether those who object to GM based on manipulation can avail themselves of the argumentative space left by the similarity argument.

Consider first the possibility that manipulation is in fact a *pro tanto* reason to object to earlier technologies as well, but that these are not all-things-considered morally problematic. One problem is that the fact that we have domesticated plants and animals has determined the future of our species to such an extreme degree that arguing that it was all-things-considered wrong of us to do so generates a version of the so-called “apology paradox”, wherein we apologise for a fact that is a necessary condition for our own existence (Thompson, 2010). Luckily, I think we can largely sidestep this issue, since most people would accept the direct intuition that manipulation is not wrong in the case of domestication and hybridization. Furthermore, the philosophers who object to GM based on manipulation tend also to accept that the manipulation done through domestication is not *pro tanto* problematic.

Next, consider possible contextual considerations that rationalize viewing manipulation as a reason to find GM problematic, but not a reason to find domestication problematic. I can think of two views. First, one underlying reason for objection to GM on the basis of manipulation is that it is risky, not in the sense that GM is likely to be harmful, but that we are not able to predict sufficiently well what effects, positive or negative, developing and using GM organisms might have. For example, Joachim Boldt has argued that the development of wholly novel organisms in synthetic biology challenges risk assessment, since organisms’ behaviour can only be known by observing them ‘in the wild’ (Boldt, 2013). Since other methods for creating organisms work by changing organisms from the outside, we

typically have a better sense of how the novel organism behaves and what effects it has in the setting in which it will be used. Second, another worry that underlies the manipulation objection is that it will lead us to treat living beings with less respect than they are due sometime in the future, since manipulation is or encourages viewing organisms as mere machines. Since we already know the effects of domestication and scientific breeding on our relationship with other living beings, this underlying worry is not relevant in the case of the comparison technologies.

Third, consider the possibility that the similarity argument employs a different sense of ‘manipulation’ than that used by the GM critics. This is in fact typically the case, at least when we are talking about philosophical critics. These typically stress that there is an importance difference in degree and/or kind between how we manipulate organisms in GM and how we do so in selective breeding and similar technologies. Borrowing a term from Beth Preston (2013), our interventions in organisms’ genotypes have become more “refined”, in the sense that we exercise more precise control and employ more detailed knowledge in doing so.<sup>7</sup> Several GM critics argue that this difference in refinement makes a moral difference. For example Keekok Lee (1999) argues that more refined interventions produce organisms that are more artefactual, since they embody human aims and intentions to a higher degree than organisms produced by less refined methods. Since artefactual entities lack a kind of value that non-artefactual (natural) entities have, this means that GM organisms are less valuable than non-GM organisms. Christopher Preston (2012) argues along the same lines as Lee, but suggests that it is only when organisms lose connection to the process of evolution – as some organisms envisaged by synthetic biologists would argu-

---

<sup>7</sup> Histories of such interventions that support this view are provided by Richard W. Bulleit (2005) for the case of animals, and by Denis J. Murphy (2007) for the case of plants. It is plausible, according to these accounts, that early domestication was entirely *unintentional*, i.e. that humans exercised no conscious control of organisms’ genotype.

ably do – that they become less valuable than other organisms. And several authors, including Lee, Cohen, Bill McKibben (2003, p. 178) and Joachim Boldt and Oliver Müller (Boldt & Müller, 2008) argue that having *too much* control is problematic, e.g. because it encourages lack of respect for other living beings, or because it is detrimental to human well-being to live in a world entirely shaped by our own intentions.

Finally, consider whether manipulation-based objections could be aimed at the conclusion that GM is conditionally acceptable. It is somewhat hard to tell, since there is a marked lack of policy recommendations and explicit normative conclusions in the writings of those who object to GM based on manipulation. But there are at least two kinds of conclusion that (as far as I can tell) are often sought, and which amount to some form of conditional acceptability. The first, exemplified by Boldt & Müller, is that GM technologies (specifically synthetic biology in Boldt & Müller's case) should be subject to some kind of ethical oversight structure – for example that an ethical code of conduct should be put in place, or that new applications should be discussed by ethicists and/or in public deliberation. The second possible conclusion, exemplified by Lee and Preston, is that natural (unmanipulated) entities have a value that manipulated entities are lacking. In effect, this means that there is a *pro tanto* reason to create or preserve a natural organism rather than to create a GM organism. This does not rationalize a general rejection of GM, but only resistance to those instances where GM organisms *replace* natural ones. Furthermore it is at least an open question how *much* extra value naturalness adds, since presumably there are also other values at play when we are deciding whether to let a natural or an unnatural organism exist.

## **5. Conclusions**

I have argued that comparisons of GM and established technologies are frequently used as part of an argument, namely the similarity argument as put forward in §2. I have used a



generic version of the argument to show that it leaves some parts of the argumentative space open. I have then used this to show to how certain common objections to GM can avoid the seeming inconsistency that the similarity argument highlights. This simultaneously shows how these objections *must* be framed if they are to avoid the inconsistency. Before ending, I would like to offer some general reflections on the usefulness of similarity arguments in the light of the above.

First, the fact that similarity arguments close off parts of the argumentative space contributes to understanding the worries that GM critics have, and thus forms the basis for a more detailed critique of their objections. For example, my analysis suggests that the belief that GM food is less healthy than 'natural' food is important, and that the notion that GM is essentially an extension of chemically intensive industrialized farming is crucial. In some cases, the position that the GM critic must carve out to avoid the inconsistency that a similarity argument alleges is vulnerable to other objections (arguably this is the case for the health-related naturalness objection). In other cases, answering the similarity argument requires formulating an objection in a more limited way than it might have been understood (arguably this is true for Lee's and Preston's version of the manipulation objection). But in neither case is a similarity argument sufficient in itself; or at least is more fruitfully used as one element in a larger argumentative strategy.

Second, there are other uses of similarity arguments, or at least similarity *claims*. In particular, they can be used as intuition generators, in the sense that they elicit the direct intuition that F is not a reason. In this case, the similarity argument may seem redundant – we could instead just directly argue that F is not a reason. However, it is sometimes useful to consider the moral relevance of a feature in a relatively familiar set of circumstances, rather than abstractly. Features described abstractly can sometimes *seem* relevant because the abstraction makes it more difficult get a clear picture of what having the feature actually

amounts to. I think this is true in the case of the widespread idea (among lay people) that all living things are valuable – which tends to be undermined by noting that we are not troubled by the murder of millions of bacteria that each of us commit daily.<sup>8</sup> Similarly, reflecting on the fact that age-old crops like corn are results of human manipulation of DNA may make vivid the intuition that manipulation is not morally dubious.

Third, as I have suggested at several places, it matters greatly for the effectiveness of similarity arguments how objections are interpreted. In many cases, objections are not stated in a detailed and rigorous manner, especially where members of the public are concerned. It seems to me that the interpretative virtues of charity and avoiding straw men should lead to a presumption that GM critics probably do not have an objection in mind that falls victim to a similarity argument.<sup>9</sup> This goes even for the weak conclusion, at least in some contexts. The weak conclusion seeks to place some explanatory burden on the GM critic. In cases where the similarity argument is used in the course of a dialogue with a critic, this is of course a fair move. But in many cases, the audience of a similarity argument is not the critic herself, but rather a third party. In such cases, the boundary between the weak and strong conclusions is blurred, especially if it is suggested that the critic would have a hard time discharging her explanatory burden. Apart from the interpretative vices that are potentially exhibited, there is also evidence that those who use so-called rebuttal analogies – of which the similarity argument may be seen as an example – are judged as being less competent and less ethical debaters by third-person observers of debates (Whaley, 1998). It may thus not be a good idea from a pragmatic point of view to use similarity arguments if the aim is to defend GM and convince readers or listeners that GM is ethically sound.

---

<sup>8</sup> Of course, biocentrists within environmental ethics have ways of arguing that our intuitions about the moral standing of microorganisms and plants does not decide the issue of their moral standing.

<sup>9</sup> Tom Douglas has suggested (in conversation) that some similarity arguments are used simply as interpretative tools, designed to bring out what the critic's view most plausibly is.

## **References**

- Baertschi, B. (2012). Defeating the Argument from Hubris. *Bioethics*, 27(8), 435-441.
- Blackburn, S. (1996). Securing the notes: Moral epistemology for the quasi-realist'. In W. Sinnott-Armstrong & M. Timmons (Eds.) *Moral Knowledge? New Readings in Moral Epistemology*. Oxford: Oxford University Press.
- Boldt, J. (2013). Life as a Technological Product: Philosophical and Ethical Aspects of Synthetic Biology. *Biological Theory*, 8, 391-401
- Boldt, J., & Müller, O. (2008). Newtons of the Leaves of Grass. *Nature Biotechnology* 26, 387-389.
- Bulleit, R.W. (2005). *Hunters, Herders and Hamburgers: The Past and Future of Human-Animal Relationships*. Columbia University Press, New York.
- Cohen, G.A. (2011). Rescuing Conservatism. In R.J. Wallace, R. Kumar & S. Freeman (Eds.) *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon* (pp. 203-230). Oxford: Oxford University Press.
- Comstock, G. (2010). Ethics and Genetically Modified Food. In F.T. Gottwald, H.W. Ingen-siep & M. Meinhard (Eds.) *Food Ethics* (pp. 49-66). Dordrecht: Springer.
- Dancy, J. (1993). *Moral Reasons*. Cambridge, MA: Blackwell.

- Deckers, J. (2005): Are Scientists Right and Non-Scientists Wrong? Reflections on Discussions of GM. *Journal of Agricultural and Environmental Ethics*, 18, 451-478
- Douglas, T., Powell, R., & Savulescu, J. (2013). Is the Creation of Artificial Life Morally Significant? *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 688-696.
- Gaskell, G., Stares, S., Allansdottir, A., Allum, N., Castro, P., Esmer, Y., ... Wagner, W. (2010). Europeans and biotechnology in 2010: Winds of change? *Report to the European Commissions Directorate-General for Research*. Retrieved from [http://ec.europa.eu/public\\_opinion/archives/ebs/ebs\\_341\\_winds\\_en.pdf](http://ec.europa.eu/public_opinion/archives/ebs/ebs_341_winds_en.pdf).
- Heyd, D. (2012). Is there anything unique in the ethics of synthetic biology? *Perspectives in Biology and Medicine*, 55(4), 55:581–589.
- Holtug, N. (2009). Creating and Patenting New Life Forms. In H. Kuhse, P. Singer P (Eds.), *A Companion to Bioethics* (2<sup>nd</sup> ed.) (pp. 235-244). Malden, MA: Wiley-Blackwell.
- Horty, J.F. (2007). Reasons as Defaults. *Philosophers' Imprint*, 7, 1–28.
- Horty, J.F. (2012) *Reasons as Defaults*. Oxford: Oxford University Press.
- 'Labels for GMO Foods Are a Bad Idea' (2013, Sep 1). [Editorial] *Scientific American*, retrieved from <http://www.scientificamerican.com/article/labels-for-gmo-foods-are-a-bad-idea/>.

- Lee, K. (1999): *The Natural and the Artefactual: The Implications of Deep Science and Deep Technology for Environmental Philosophy*. Lanham, MD: Lexington Books.
- Marris, C., Wynne, B., Simmons, P., Weldon, S. (2001). Public Perceptions of Agricultural Biotechnologies in Europe. *Final Report of the PABE research project*. Retrieved from [www.csec.lancs.ac.uk/archive/pabe/docs/pabe\\_finalreport.pdf](http://www.csec.lancs.ac.uk/archive/pabe/docs/pabe_finalreport.pdf).
- McKibben, B. (2003). *The End of Nature* (revised ed.). London: Random House.
- Mielby, H., Sandøe, P., Lassen, J. (2013). Multiple aspects of unnaturalness: Are cisgenic crops perceived as being more natural and more acceptable than transgenic crops? *Agriculture and Human Values*, 30, 471-480.
- Murphy, D.J. (2007). *People, Plants and Genes: The Story of Crops and Humanity*. Oxford: Oxford University Press.
- Nozick, R. (1974). *Anarchy, State and Utopia*. New York, NY: Basic Books.
- Preston, B. (2013). Synthetic Biology as Red Herring. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 649-659.
- Preston, C.J. (2008). Synthetic Biology: Drawing a Line in Darwin's Sand. *Environmental Values*, 17(1), 23-39

- Pugh, J. (2015a, December 2). The Ethics of Genetically Modified Mosquitoes and Gene-Drive Technology. [Web log post]. Retrieved from <http://blog.practicaethics.ox.ac.uk/2015/12/the-ethics-of-genetically-modified-mosquitoes-and-gene-drive-technology/>.
- Pugh J (2015b, December 2). Playing God with mosquitoes? We humans have loftier aims. *The Conversation*. Retrieved from <https://theconversation.com/playing-god-with-mosquitoes-we-humans-have-loftier-aims-51362>.
- Rozin, P. (2005). The Meaning of “Natural”: Process More Important Than Content. *Psychological Science*, 16(8), 652-658.
- Rozin, P., Spranca, M., Krieger, Z., Neuhaus, R., Surillo, D., Swerdlin, A. & Wood, K. (2004). Preference for natural: instrumental and ideational/moral motivations, and the contrast between foods and medicines. *Appetite*, 43, 147-154
- Rozin, P., Fischler, C., Shields-Argelès, C. (2009). Additivity Dominance: Additives are more potent and more often lexicalized across languages than “subtractives”. *Judgment and Decision Making*, 4(5), 475-478
- Sandler, R.L. (2007). *Character and Environment: A Virtue-Oriented Approach to Environmental Ethics*. New York, NY: Columbia University Press.
- Swierstra, T., & Rip, A. (2007). Nano-Ethics as NEST-Ethics: Patterns of Moral Argumentation About New and Emerging Science and Technology. *NanoEthics*, 1(1), 3-20.

Thompson, J. (2010). The Apology Paradox. *The Philosophical Quarterly*, 50, 470-475.

Thompson, P.B. (2003). Unnatural Farming and the Debate over Genetic Manipulation. In  
V.V. Gehring (Ed.). *Genetic Prospects: Essays on Biotechnology, Ethics and Public Policy*.  
Lanham, MD: Rowman and Littlefield.

Thompson, R.P. (2011). *Agro-technology: A Philosophical Introduction*. Cambridge: Cambridge  
University Press.

Whaley, B.B. (1998) Evaluations of Rebuttal Analogy Users: Ethical and Competence Con-  
siderations. *Argumentation*, 12, 351-365





## Article 3 Rationality, Thresholds and the Precautionary Principle

### 1. Introduction

Many of our activities simultaneously offer the possibility of great benefit and the risk of serious harms. Technological developments allow us to improve health, happiness and environmental sustainability, but may also cause great losses to those same values. These are *risky* things. We must figure out how we are to secure the benefits without incurring the harms. And if we cannot do that, we must figure out what harms we can risk in order to keep the chance of the benefits. Such choices, where we must decide which policy to implement when each policy may lead to several different outcomes, are what I will call 'risky choices'. The precautionary principle (PP) is supposed to help us in making risky choices. At the most general level, it asks us to act to eradicate risks of severe harm, even if we must pay the price of forgoing possibly significant benefits.

Many have argued that PP is a poor guide. The categorical demand that severe risks be eradicated has seemed to many *irrational*. In this paper, I will investigate what this supposed irrationality might amount to. I suggest that there are three intuitively strong objections to PP that all make the claim that guiding our choices by PP is irrational. I will argue that these objections can be answered, at least when PP is interpreted in a suitably moderate way. The challenge I present goes beyond the much-discussed claim that PP is *incoherent* (Sunstein, 2005). By defending PP from stronger rationality objections, I hope to strengthen the case for PP as an element in sound risk management. In addition, the limits within which a defensible PP must keep itself are made clearer.

## **2. The precautionary principle**

The most famous formulation of PP is probably the Rio Declaration of 1992, which defines PP as follows: “Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation” (UNEP, 1992, §15). As several authors have shown, the basic schematic structure of the Rio Declaration can be found in most, if not all, other formulations of PP in legal and policy documents and in the academic literature (Manson, 2002; Sandin, 1999; Trouwborst, 2006, Ch. 3). That structure consists in the identification of some *harm* that an activity may cause (“threats of serious or irreversible damage”), some degree of *uncertainty* as to whether the activity will cause the harm (“lack of full scientific certainty”) and a precautionary action taken to avoid the harm (“cost-effective measures to prevent environmental degradation”). I take it to be as uncontroversial as is reasonably possible that this basic structure is definitive of PP.

There are two possible sources of confusion in this definition of PP as a relationship between harm, uncertainty and action, namely ‘uncertainty’ and ‘action’. Uncertainty is a feature of our knowledge of the relationship between some activity or policy (e.g. the emission of CO<sub>2</sub>) and some harmful outcome (e.g. global warming). But it is a *negative* feature of this knowledge. It defines a type of knowledge, namely (full scientific) certainty, as *not necessary*. This tells us very little, and in particular leaves it open what kind of knowledge is *sufficient* and/or *necessary*. Concerning action, the confusion runs in the other direction. It sounds as if there is a *positive* relationship between the harm and some action. But in fact no *particular* kind of action is typically named. At best, some desired or required *features* of the action are mentioned – in the case of the Rio Declaration, that it is “cost-effective” and seeks “to prevent environmental degradation”.

I therefore suggest that we substitute ‘knowledge’ for ‘uncertainty’ and define the relationship to action negatively, such that PP prohibits certain policies with respect to the possibly harm-producing activity.<sup>1</sup> Consequently, the basic PP schema is this: If we have some degree of *knowledge* that a policy will cause some *harm*, then that policy is prohibited. This leaves only *two* variables, namely *harm* variable (H) and the *knowledge* variable (K). These function as thresholds, such that any policy that we have knowledge at least as good as K will lead to harm at least as bad as H is prohibited. Using H and K as thresholds require that we are able to rank both outcomes and types of knowledge. I will assume here that this can be done, at least in principle – i.e. that it makes sense to rank one outcome as better or worse than another, or that we have at least as good reason to believe one thing than another. With respect to H, I will denote the overall goodness of an outcome its *utility*. With respect to K, I will follow the standard Bayesian idea that the strength of our reasons to believe some proposition and the likelihood of that proposition being true are at root identical. Consequently, instead of talking about how good reason we have to believe that a policy will lead to an outcome, I will often talk of how likely that outcome is.

## 2.1 Coherence

PP, as defined here, thus prohibits any policy that triggers H and K thresholds, and allows all other policies. In any concrete application, H and K must be set to specific levels. Following Steel, I will call an instance of PP with fixed H and K thresholds a “version” of PP (Steel, 2014, p. 27). Since PP only prohibits policies that trigger H and K, and allows all others, it does not select a unique policy in all circumstances (only when just one policy is allowed).

---

<sup>1</sup> My account here follows, and is highly indebted to Daniel Steel’s account (2014). However, Steel does not consistently use the language of prohibition, partly (I think) because it makes the *objection* to PP that he wants to reject – the incoherence objection – hard to even state.

Consider the example of so-called gain-of-function research on the H5N1 influenza virus. In gain-of-function research, pathogens are modified in ways that make them more dangerous than they currently are in order to study the mutations that may make them more dangerous in the wild. In the H5N1 case, a strain of influenza virus with extremely high mortality rates (of over 50%) was modified in a way that probably makes it transmissible between human beings (which it had not previously been). An H5N1 pandemic is estimated to kill anything between 2 million and 1.4 billion people (Rozell, 2015). If conducted in laboratories of the required safety level – biosafety level 3 (BSL-3), the risk that enhanced H5N1 virus escapes and causes a pandemic has been estimated at between 1/1.000 and 1/10.000 per laboratory year (Lipstich & Ingleby, 2015). Suppose that the version of PP we are working with sets H and K such that these fatality and probability ranges meet the thresholds. Then the policy of conducting H5N1 gain-of-function research in BSL-3 laboratories is prohibited by PP. However, no other policy recommended, since we do not yet know if any such policy is itself allowed by PP.

Consider for example a total ban on H5N1 gain-of-function research, which some might advocate on the basis of the version of PP considered. Is this policy allowed by PP? That will depend on the likelihood that a ban will cause a similar number of deaths (or more precisely that a ban will cause the failure to prevent a similar number of deaths). Since the research would be extremely valuable in case the H5N1 virus naturally mutates to become human-to-human transmissible, and since it is likely that the virus will do exactly that at some point, it is very plausible that a ban would meet the H and K thresholds and thus be prohibited. But there is likely another policy that would be allowed, namely one in which H5N1 gain-of-function research is conducted in safer laboratories than BSL-3. If this policy sufficiently lowers the likelihood that the virus escapes and causes a pandemic – such that it is below the K threshold – then it will be allowed.

This example illustrates an important virtue of the account of the PP that I defend here, namely that it avoids the incoherence objection. According to the incoherence objection, the PP prohibits every policy, and therefore also prohibits the very policy that it is used to justify (Sunstein, 2005, Ch. 1). There are two aspects to this objection. The first is that PP is *paralyzing* because it prohibits every policy that may cause some harm. Since every policy may cause some harm – even the safer research policy on H5N1 might weaken our response to a naturally-occurring pandemic somewhat – PP will prohibit every policy. The use of thresholds solves this problem, since even if all policies have *some* risks, the specific risks of any given policy can fall below either H or K (or both), as is the case for the safer research policy. The second aspect of the incoherence objection is that PP is inconsistent, because the policies it recommends sometimes create risks that are at least as bad as the policy that originally demanded precaution. But in such cases, PP as I have defined it does not recommend the equally risky policy – it prohibits it, just as it prohibited the original policy. Daniel Steel names this feature of PP the consistency principle, which he defines as the principle “that a precaution should not be precluded by the same version of PP used to justify it” (Steel, 2014, p. 28). The policy of banning H5N1 gain-of-function research altogether fails the consistency test, since it was justified with reference to a version that prohibits it.

## 2.2 Proportionality

Though PP does not recommend any unique action, a policy will have to meet further requirements than simply being allowed by the relevant version of PP. As we have seen, the Rio Declaration demands that precautionary measures are “cost-effective”. The European Commission requires policies to be “proportional to the chosen level of protection”, consistent with established legal principles such as non-discrimination, and open to review as

new research improves the knowledge base (European Commission, 2000, p. 3). The concept of proportionality is especially important here. It is a general principle in European Union law, and limits legitimate legal action to what is necessary in order to achieve the goal laid down in the relevant legislation.<sup>2</sup> An especially important aspect of this requirement is that the policy chosen should be the least restrictive one available (Harbo, 2010; Trouwborst, 2006, p. 153). Steel (2014, pp. 26-30) incorporates a proportionality requirement into his account of PP, consisting of the two sub-requirements of consistency and efficiency. I have already described the effect of consistency above. Efficiency is defined as the demand that “if more than one precaution can be consistently recommended by the version of PP being used [i.e. are allowed by the version of PP], then those with lower costs should be preferred” (Steel, 2014, p. 29). Costs should here be taken to include not only out-of-pocket costs, but also the foregone benefits from limiting technological innovation.

It is debatable whether we should see the proportionality requirement as *part of* PP, or whether it is better seen as an exogenous principle that applies when PP as such has nothing to say. I tend toward the latter, since proportionality is a *general* principle that applies to other things besides PP, and since its justification is not particularly closely related to the concept of precaution. On the other hand, proportionality is clearly not optional, but an integral part of the use of PP in legal- and policy practice. Consequently, proportionality considerations can legitimately be used to argue that a certain policy is, or is not, ultimately justified with reference to PP.

To illustrate the effect of proportionality, consider example of synthetic biology. Synthetic biology enables the design and construction of biological systems for useful purposes, such as the production of valuable substances (e.g. biofuels, oils, flavourings and pharma-

---

<sup>2</sup> See Article 5, §4 of the Treaty of the European Union, available at <http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A12012M%2FTXT>

ceuticals), monitoring and cleaning up environmental pollutants, targeted medical interventions and more. A consortium of civil society groups have called for a moratorium on “the release and commercial use of synthetic organisms”, and justified that call by reference to PP (FOE, ICTA & ETC Group, 2012, p. 3). FOE et al. do not explicitly define H and K thresholds. They merely argue that PP is triggered in the synthetic biology case since it has “potentially far-reaching and irreversible impacts”, and that “the risks of the technology are inherently unpredictable”. The moratorium is supposed to stand until several requirements have been met, including (1) the development of a research agenda based on the public interest, (2) a full consideration of alternatives have been undertaken, (3) a full assessment of health, environmental and socioeconomic impact has been undertaken, and (4) oversight and security mechanisms have been developed.

Given the vagueness of the features that trigger PP according to FOE et al., it is possible that they rely on unjustifiably low thresholds (e.g. that they implicitly require full certainty that no harm will occur). But let us suppose that synthetic biology, if unregulated, poses risks to health and environment that meet reasonable H and K thresholds, and that the moratorium would be allowed by the relevant version of PP. The ultimate justifiability of the moratorium is not settled by these facts alone, since there may be other policies that are also allowed, and that are more efficient (in Steel’s terminology). For example, a moratorium without requirements (1), (2) and the socioeconomic part of (3) would prevent risks to health and environment just as well, and would be less restrictive. Similarly, a policy that restricts only those particular applications of synthetic biology that have a reasonable likelihood of leading to harm, and not innocuous ones, would be preferable on proportionality grounds. Furthermore, proportionality will be important when choosing the appropriate oversight and security mechanisms, as per demand (4) of FOE et al. It is unlikely that banning synthetic biology applications will be necessary, since less restrictive policies (such as

extensive pre-release testing) will in most cases be sufficient. On the other hand, PP *may* justify a ban of some applications of synthetic biology, if there is no other policy that would reduce risks to an acceptable level.

### **3. The rationality objection(s)**

As we have seen, PP as defined in the previous section avoids the incoherence objection, since it neither paralyzes policy-making by prohibiting all risk-creating activities, nor recommends precautionary measures that are equally likely to lead to equally bad outcomes as the original policy. The use of thresholds of harm and knowledge ensures this result by allowing that some risks are too small to warrant precaution, and by allowing only policies that are consistent with the version of PP set by the relevant H and K thresholds.

However, the incoherence objection is not the only objection to PP – it is not even the only objection that seeks to show that PP is *irrational*. In the eyes of many, we already have an adequate theory of rationality that applies to risky choice, namely *expected utility theory* (EUT). According to EUT we should choose the policy that *maximizes* expected utility, which is defined as the sum of the utilities of each possible outcome of a policy multiplied by the probability that the outcome will obtain. In this section, I will show how the ideas used to argue that EUT is the correct normative of risky choice can also be used to mount objections to PP. Interestingly, these objections are objections to the *very use* of thresholds of knowledge and/or harm (or more generally, value) in deciding what to do in situations of risk. So the same feature of PP that allows it to avoid one rationality-based objection, namely the incoherence objection, may generate other rationality based objections. I identify three objections of this sort: (i) The use of PP will lead to sub-optimal outcomes in the long run; (ii) PP treats some relevant differences in the utility and/or likelihood of outcomes as irrele-



vant; (iii) Thresholds arbitrarily endow specific utilities and/or likelihoods with special importance.

### **3.1 The long-run objection**

A main argument for using EUT as the normative theory of risky choice is that it ensures that we are as well off as possible in the long run. Any divergence from EUT, including the use of thresholds, therefore predictably leads to long-run outcomes that are worse than what we could have achieved by using EUT. EUT, recall, calls for the maximization of expected utility, i.e. that we choose the policy that has the largest sum of probability-weighted utility. Due to the law of large numbers, maximizing expected utility in each individual choice will almost certainly result in maximizing *actual* utility in the long run. According to the law of large numbers, the actual value of a random variable will converge to its statistical expectation in the long run. In risky choice, the actual utility achieved is a random variable, since it depends on which of the possible outcomes of a policy is realized. Since EUT maximizes expected utility, it ensures that the expected value to which actual utility converges in the long run is the largest one possible. And since the expected utility of any other strategy for risky choice, including PP, is lower than the maximum, we are statistically certain to end up with less than the maximum actual utility in the long run by using PP. Since utility is a measure of everything we care about, PP is irrational: It is a decision principle that predictably fails to best achieve what we want to achieve.

### **3.2 The neglect objection**

While the long-run objection concerns a choice strategy – that is, the use of a principle like expected utility maximization or PP generally over many risky choices – the second objection concern single choices. We can make at least some judgments about how strong reasons

we have for choosing a policy in an individual risky choice situation. Call this the *choiceworthiness* of policies. According to the second objection, there are some changes in the intuitive relative choiceworthiness of policies that PP neglects.

There are two possible kinds of change that can be made to a policy, namely a change to the relative<sup>3</sup> likelihood of outcomes and changes to the utilities of outcomes. First, consider changes to likelihood. Suppose we have two outcomes, a good and a bad one. Intuitively, lowering the relative likelihood of the bad outcome – and increasing the likelihood of a good outcome – improves the choiceworthiness of a policy. Similarly, increasing the relative likelihood of the bad outcome worsens the choiceworthiness of a policy. More generally *moving* likelihood from a worse to a better outcome improves choiceworthiness and moving likelihood from a better to a worse outcome worsens choiceworthiness. Second, consider changes to utility. The principle here is even clearer, since utility is not relative in the same way that likelihood is. Quite simply, a policy's choiceworthiness is improved if the utility of one of its possible outcomes is increased, and its choiceworthiness is worsened if the utility of an outcome is decreased.

PP does not in all cases respect changes of these kinds of change. There are two kinds of problem. The first is that PP treats two policies as equally choiceworthy although one dominates the other. A policy dominates another if it is more choiceworthy than another in any of the ways described above. Consider two policies, A and B, that have the same two possible outcomes, Good and Bad. Suppose Bad is sufficiently bad to meet the H threshold, and that the likelihood of Bad is lower in A than in B. A thus dominates B in terms of likeli-

---

<sup>3</sup> I assume that the likelihood of one outcome is relative to the likelihood of other outcomes, since the laws of probability demand that the likelihood of all possible outcomes sum to 1; decreasing the likelihood of one outcome therefore necessarily increases the likelihood of another outcome (although we may not know which other outcome's likelihood is increased, and although the likelihoods may not be known with much precision).

hood. As long as the likelihood of Bad in both A and B is above the K threshold, both policies will be equally prohibited by PP.

Next consider two other policies, C and D. Each has two possible outcomes, Likely and Unlikely. The likelihoods are identical for C and D. The utility of unlikely is very low in both cases, but the utility of Likely-C is higher than Likely-D. C therefore dominates D. But supposing that Unlikely meets the K threshold, both C and D will be prohibited, and the dominance of C will not show itself in the recommendations of PP. PP thus neglects dominance, both in terms of likelihood and in terms of utility.

The other problem has to do with trade-offs. Suppose we have two policies, E and F, that each has a good and a bad outcome. Suppose Bad-F is sufficiently bad to meet the H threshold, but that Bad-E is not. By improving the utility of Good-F, we make F more choiceworthy. But as long as F is prohibited by PP, we can make Good-F *arbitrarily* better without ever making F more choiceworthy than E. It thus seems that the use of thresholds implies that a prohibited policy's choiceworthiness can be improved by an infinite amount, and yet it can never be more choiceworthy than an allowed policy.

### **3.3 The arbitrariness objection**

The third and final objection is that the PP's use of thresholds of utility and likelihood arbitrarily places a large importance on very specific levels of utility and likelihood. Why should exactly *these* levels determine the difference between prohibition and allowance? The objection comes in three different strengths. The weak version merely demands that we should be prepared to justify the use of the specific relevant threshold levels when arguing for the use of some version of PP in some individual case. Justification may be wholly case-specific, and there are no general criteria for the kinds of values H and K must take for a version of PP to be warranted. The medium-strength objection demands exactly such gen-

eral criteria. In other words, it argues that PP must include a list of the kinds of values and likelihoods that can figure in a justifiable version of PP, and that we can give reasons why these kinds of values and likelihoods should be accorded the importance that serving as thresholds endow them with. Finally, the strong version of the objection claims that no reason *can* be given for endowing any values and likelihoods with this sort of importance.

According to the weak and medium-strength objections, PP is irrational because it does not support important judgments with reasons. The strong version of the arbitrariness objection views PP as irrational because the very idea of thresholds violates some basic norms that our normative thinking must recognize. The kind of *categorical* distinctions between values and/or likelihoods that thresholds require are simply not part of any justifiable normative theory, moral or prudential.

### **3.4 Relations between the objections**

Although the three objections focus on different issues, they are clearly related, and they mutually support each other in several ways. The long-run objection strengthens the intuition that changes in choiceworthiness should matter. Since the intuitive changes in choiceworthiness are also changes in expected utility, the long-run perspective makes vivid the costs of a strategy that allows the choice of less-choiceworthy policies. Similarly, the long-run objection highlights a danger in case-specific justification of thresholds, since the cumulative effects of choices that seem acceptable in isolation may be large and bad. Finally, there is a close relationship between the strong arbitrariness objection and the trade-off version of the neglect objection. The idea behind the trade-off version of the neglect objection is that there must be *some* increase to the utility of a good outcome that compensates for the badness of a bad outcome. Martin Peterson (2008, p. 115; 2009, p. 76) defends this “trade-off

principle” as an axiom of rationality. In another paper (2006, p. 599) he argues for an analogous principle with respect to likelihoods.

The three objections all challenge the use of thresholds of likelihood and of value in risky choice, although they do it from slightly different angles. Furthermore, as noted, the weaknesses of PP that they stress are the flipside of important virtues of the theory the objections are grounded on, namely EUT. The overall challenge they pose to PP is thus this: How, if at all, can we justify the use of a decision procedure that uses thresholds of utility and likelihood? Such a procedure is inferior to EUT in the long run, and it seems not to respect the relevance of value and likelihood changes for the choiceworthiness of policies. So why not just use EUT, which is known to produce optimal outcomes in the long run, and does not make use of counterintuitive categorical distinctions?

#### **4. Answering the objections**

My answer to the challenge has two parts. First, I show that the use of categorical distinctions in value is more plausible than one might suspect, and hence that EUT may not be the correct normative theory. Second, I argue that even if we assume that EUT is the correct normative theory, we still might be well advised to use PP, since it does better than the non-ideal implementations of EUT that are its realistic competitors.

##### **4.1 Categorical distinctions in value**

Several of the rationality objections either assume or try to argue that categorical distinctions in value, such as PP’s use of the H threshold, cannot be justified. The strong neglect and arbitrariness objections state this directly, while the weaker versions merely challenges defenders of PP to justify the thresholds, or to provide an argument for why the seemingly relevant changes in relative choiceworthiness cannot make a prohibited theory more

choiceworthy than an allowed policy. But, as I argue in this section, several plausible normative theories do allow for categorical distinctions.

Arguably the most likely source of categorical distinctions is the view that utilities are defined *relationally*, i.e. with reference to the other outcomes that are at play in a given choice. A familiar idea in much (especially non-consequentialist) ethical theory is that certain trade-offs are off-limits, e.g. that human lives may not be traded off for mere economic gain. In the case of risky choice the claim must be that a specific type of bads may not be *risks* in order to gain a chance of a specific type of good. It seems intuitively plausible, for example, that we ought not risk our lives for minor economic gains – say \$10 – no matter how low the likelihood of losing is; and, to return to the synthetic biology example, that we ought not risk human extinction for the sake of lower prices on agricultural products. In the European Commission's guidelines on PP, it is repeatedly stressed that the protection of public health should take precedence over economic considerations (e.g. European Commission, 2000, p. 4, 19). The implication of a relational theory is that changes of likelihood or utility of the kinds stressed by the neglect objection *should* never make a difference to the ranking of policies – i.e. that no prohibited policy should ever be chosen over an allowed policy.

The other general strategy for justifying categorical distinctions in value takes it to be the case that some bads are *intrinsically* categorically worse than most others. In the context of PP, two kinds of outcome are often highlighted, namely *irreversible* and *catastrophic* ones. It is not difficult to see why catastrophes have extremely low utility. It is less obvious for irreversible outcomes, since an irreversible outcome could be good, or bad but trivial. The kind of irreversible outcome that could generate categorical differences in utility occurs when some valuable thing, or some bearer of value, goes out of existence. Examples include the extinction of species, including humanity, the destruction of ecosystems or historical

artefacts, the bankruptcy of a business and the death of a human being. The general idea behind seeing irreversible outcomes as categorically worse than reversible ones is that valuable things, or bearers of value, generate utility at each point in time in which they exist. Valuable things can do so by having 'existence value' or because contemplating them generates valuable mental states; bearers of value can do so by being necessary for the occurrence of valuable features of states of affairs, such as welfare.

An especially strong version of the irreversibility problem concerns cases where what is lost is *everything* that generates value. This is known as the *ruin* problem. Importantly, ruin undermines the relevance long-run considerations, since ruin prevents the long run from ever happening.<sup>4</sup> From the point of view of an investor, for example, bankruptcy literally wipes out everything. Most cases of ruin depend on the framing of decisions; for example, my death is ruinous when considering how choices affect value-for-me, but of course does not amount to the end of all value. Perhaps the extinction of humanity, or of all life, would be ruinous in a global framing as well, if the very existence of value depends on human consciousness, sentience, or teleological striving. At least for all we know it could be, if no such things exist elsewhere in the universe, and will not evolve there either.

However, I am sceptical that either catastrophes or irreversible loss of valuable things are genuinely categorically worse than other types of outcome. Catastrophes may be thought of as merely very, very bad. And the value loss of irreversible outcomes will also be bounded, if the alternative to destruction now is not existence forever, but merely existence for some time. The orthodox theory of the badness of death, for example, says that it is equal to the good life that the person would have experienced had she not died now, but at the later time where she would have otherwise died (see McMahan, 2002, pp. 103-107). If

---

<sup>4</sup> The long run is, of course, not literally an event. Rather, long-run value is a limit value as time (or number of choices made) approaches infinity.

other valuable things also have a limited existence, which they plausibly do, then the negative utility of destroying them will similarly be bounded (although potentially enormous). Furthermore, it is noteworthy that many outcomes are not irreversible as such, but merely 'sticky' in the sense that some source of utility is knocked out for a considerable length of time, but not forever. For example, pollution and some ecosystem collapses are sticky rather than irreversible. The value loss of such outcomes will be potentially much lower than a similar but genuinely irreversible outcome.

Two considerations nevertheless justify treating irreversible outcomes as categorically worse than others. First, such outcomes may not be capable of being outweighed by corresponding good outcomes, since there are no good outcomes that have a positive utility of the same magnitude as the negative utility of the irreversible or catastrophic outcome. This is especially plausible when irreversible bads are compared with goods that only generate utility at the current time. For example, no improvement in my welfare level seems comparable to the loss of, say, 50 years of 'normal' good life. More generally, we might suppose that utility is bounded upwards, i.e. that there is a maximum goodness that an outcome can have. In that case, the magnitude of badness of very bad outcomes may in effect be categorically greater than the magnitude of goodness of even the best possible outcome. An important caveat is that these considerations do not apply if goodness can include 'opportunity goodness', i.e. if we define the value of *preventing* a very bad outcome as 'good'. The notion that opportunity goodness of preventing catastrophes is categorically smaller than the ordinary badness of the catastrophe seems to require a very strong form of the doing/allowing distinction (or a related distinction) that is *prima facie* implausible.

Second, categorical distinctions can be pragmatically justified. The difference in utility is judged to be so large that it makes sense to treat two outcomes as being of categorically different value, although *really* they are not. The advantage to this move is that we can ig-



nore the details of precisely how much worse a catastrophe of some kind is than some other, more 'normal' outcome. This line of argument has been applied to human extinction (Parfit 1984, pp. 453-454), global warming (Shue, 2015) and nuclear weapons proliferation (Goodin, 1985; McMahan, 1986). In the context of PP, Alan Randall (2012, pp. 110-121) has argued that precaution is warranted when a bad possible outcome is *disproportionate* to the good outcome(s) that are also possible. In slogan form, "Don't risk great harm in pursuit of moderate benefit" (Randall, 2012, p. 112). As this phrase suggests, Randall's theory looks more like the relational theories discussed above. The important difference from the intrinsic theories is that Randall allows for precaution in cases where the bad outcome is not *catastrophically* bad, but merely bad relative to the good outcomes involved.

I have not here attempted to argue that categorical distinctions are ultimately part of the correct normative theory. My aim has merely been to show that plausible normative theories do allow for categorical distinctions, and hence that it is at least possible that the categorical distinctions used by PP are justified by the correct normative theory. This possibility weakens the rationality objections, since acting in accordance with the correct normative theory can hardly be irrational.

## **4.2 Precaution versus real-world utility maximization**

Suppose that it turns out that EUT is the correct normative theory of risky choice. PP might still be justified pragmatically, as the best available decision procedure to follow in the real world. This is exactly what I will now argue that it is. I argue first that the ideal EUT that inputs the correct utilities is not the same as non-ideal EUT (that is, real-world utility decision procedures that directly aim to maximize utility, such as cost-benefit analysis). Second, I argue that PP may not diverge from ideal EUT in its recommendations of policies. Third,

and finally, I argue that PP is superior to non-ideal EUT given the circumstances in which we must choose.

#### ***4.2.1 Ideal and non-ideal EUT***

Ideal EUT is the version of EUT in which all the utilities we use as inputs are the right ones. The project of finding the right utilities faces two problems, namely that it requires a large number of very hard-to-make judgments, and that utilities may not be independent of moral and prudential judgments about what risky choices to make.

Consider the first problem. In order to arrive at precise measures of the utility of outcomes, we must be able to assess the relative values of all the goods and bads that contribute to utility. Just to cover ordinary environmental and public health policy, this would require assessing the relative values of technological innovations in a number of fields, purely financial losses and gains, the preservation of ecosystems, the welfare of animals, detrimental health effects on human beings and the loss of human life. Clearly this is a monumental task. Markets are the only example in which such a number of relative values are determined precisely, in the form of the relative prices of goods. But this is arguably only possible because the relative values that are embodied in the price system do not correspond to any judgments of relative value, but are rather the aggregated effects of many small (more or less qualified) judgments. It is doubtful whether the aggregation of this enormous number of small judgments can be mimicked by any non-market system, and also whether such aggregated relative values are normatively compelling.

The second, and arguably harder, problem is that determination of utilities is not independent of judgments involving risk. EUT requires cardinal utility numbers, rather than merely ordinal ones. Ordinal numbers (which I have assumed it is possible to assign) give us only a ranking, such that a higher number signifies that the outcome is better. Car-

dinal numbers also provide relative magnitudes, such that an outcome with utility 6 is twice as good as an outcome with utility 3. Calculating expected utility requires cardinal utilities. In orthodox axiomatic versions of EUT, cardinal utilities are arrived at by eliciting preferences over risky choices. It is therefore trivially correct that we should maximize expected utility, since the cardinal utilities are defined as those that make maximizing expected utility correct. If relative utilities are supposed to be morally defensible, the orthodox method would therefore require us to have a full and defensible moral theory of risky choice before we can arrive at the utilities that should serve as input to ideal EUT.

But suppose we reject the orthodox way of determining cardinal utility, and assume that we can assess such utilities independently of judgments about what risky choices are justified. Ideal EUT will need to accommodate the everyday observation that people tend to be risk averse in the sense that they prefer not to take gambles with small expected gains in terms of money (or other goods) if the possible losses are large – e.g. a coin toss where heads yield \$10,000 and tails yields \$-9,000. The standard way in which EUT accommodates common sense risk aversion is by assuming that the marginal utility of money is diminishing. That is, the utility derived from gaining an extra  $X$  dollars is smaller the more money you already have. It makes sense to prefer the status quo to the coin toss because the \$10,000 potentially gained is worth less, in terms of utility, than the \$9,000 potentially lost. Even though the gamble has an expected *gain* in terms of money, it has an expected *loss* in terms of utility. The same potentially goes for all other bads and goods that contribute to utility.

Diminishing marginal utility makes the task of determining the utility of an outcome much harder, because we cannot merely determine the amount of the various goods and bads and then multiply those by their respective value. The reason is that there is no fixed value for a unit of a good – the value of a unit of a good is dependent on how much of the good we already have. Diminishing marginal utility is merely one example of *holism* in val-

ue, i.e. of how the value of some thing depends on the presence or absence of other valuable (or disvaluable) things (see e.g. Brown, 2007; Schroeder, 2011). The notion of diminishing marginal utility generates a further problem for EUT, since it implies that there is a *de facto* upper boundary of utility: As the current level of utility grows, the extra amount of a good (e.g. wealth or well-being) that needs to be added to achieve an extra unit of utility approaches infinity. It is of course debatable where this *de facto* boundary is. But the mere fact that it is there supports the pragmatic use of categorical distinctions, as argued in §4.1.

Determining what the utilities in ideal EUT are is thus a very difficult, perhaps impossible task. Any non-ideal EUT procedure, such as cost-benefit analysis, thus at best approximates ideal EUT. Such non-ideal procedures therefore do not straightforwardly inherit the normative status of ideal EUT (which I assume here to be the correct normative theory), and we cannot assume that they recommend the same policies that ideal EUT would.

#### **4.2.2 Ideal EUT and PP**

The rationality objections assume that PP is a more or less *radical* departure from ideal EUT. I will now argue that this assumption is not justified. Since it is not clear what ideal EUT recommends, which rules out a direct comparison of recommendations, I can only show that PP does not (contrary to popular belief) possess certain features that would plausible bring it into disagreement with ideal EUT.

First, PP is often assumed to be a principle that only justifies policies that are maximally safe. But PP primarily prohibits certain policies, namely those that have a likelihood greater than K of leading to an outcome worse than H. As I discussed above, it is open for debate whether the principles used to select among the policies allowed by a version of PP should be seen as part of PP as well. I suggested that they should not. At any rate, the important thing is that these principles are not precautionary in any sense. They may be risk-

neutral, or at most risk-averse in the ordinary way that I have argued ideal EUT should be able to accommodate.

Second, PP could diverge quite a lot from ideal EUT if the H and K thresholds could be set at any level (cf. the arbitrariness objection). But they cannot. The most fundamental point is that thresholds must be justified independently; setting thresholds to justify one's preferred policy is a misuse of PP. In practice, any application of PP would occur in institutional settings where H and K would need to be justified, e.g. in a court of law, in administrative practice or in ordinary political deliberation (Trouwborst, 2006, pp. 99-111). Furthermore, PP is already circumscribed by certain criteria for what kind of outcomes and likelihoods can serve as thresholds, or if not it can easily be supplemented by such criteria. With respect to H, it is important the harm is genuinely "special bad", in Stephen John's phrase (John, 2007). The considerations described in §4.1 precisely aim to show that special bads are involved, whether they are intrinsically or only relationally special.

With respect to K, PP will normally require at least the *use* of the best available scientific data. If such data clearly indicates that no harm is at all likely, then PP does not come into function. For example, the European Commission requires that "preliminary objective scientific evaluation indicates that there are reasonable grounds for concern" (European Commission, 2000, p. 2). And contrary to what critics of PP often suppose, the mere *possibility* of harm is not sufficient. Critics are not solely responsible for this, since some of those who use PP to justify policies *rhetorically* also frequently imply that mere possibility is sufficient. But PP as it is realized in practice demands more.

#### **4.2.3 PP versus non-ideal EUT**

I have now argued, in effect, that there is no strong reason to think that non-ideal EUT comes much closer to ideal EUT than PP does. Still, given that non-ideal EUT at least at-

tempts to estimate what ideal EUT would recommend, it might be reasonable to think that we should prefer using non-ideal EUT as our decision procedure for risky choice. But as I will now argue, there are in fact reasons to prefer PP as a decision procedure. The general reason for this is that the circumstances in which our choices must be made favour using PP, rather than non-ideal EUT.

The first way in which circumstances militate against non-ideal EUT is well known in discussions of risky choice. Just like EUT requires cardinal utilities, it also requires quantitative information about likelihoods, i.e. that probabilities can be assigned to each outcome. PP is often defended as a principle that is supposed to apply in exactly those cases where it is impossible to assign precise probabilities (although the PP I have defended is not in principle limited to such cases). EUT on the other hand has (strictly speaking) nothing to say about these cases. If EUT is to be action guiding in a case where probabilities are missing, it must be supplemented by an extra decision principle (e.g. that regulation is only justified if an EUT can show that benefits justify costs). Such extra principles are not automatically justified by the same reasons that justify EUT, and there is a distinct risk that they are applied unconsciously and therefore without being properly thought through.

The second problem for EUT is a generalization of the first. Successful implementation of EUT does not merely require that we can assign probabilities and cardinal utilities to outcomes, but also that this information is *precise* and of *good quality*. Information is precise when we have a single-number value for probability or utility, rather than an interval of possible values, or a rough estimate. Information is of good quality if it is based on solid evidence of the kind relevant to probabilities and utilities, respectively. In most cases, our information will have at most one of these properties.

Consider for example the case of H5N1 gain-of-function research. In that case we have an estimate based on relatively good evidence of the likelihood that the virus will es-

cape from a BSL-3 laboratory. But that estimate ranges from 1/1,000 to 1/10,000, which generates large differences in the expected utility of the policy if the (positive or negative) utility associated with that outcome is large. In the H5N1 case, it is: Namely expected fatalities of between 2 million and 1.4 billion. The uncertainty in the probability estimates thus generates a difference in the expected value of the BSL-3 policy of between 18,000 and 1.26 million lives lost. Although there are generalizations of EUT that allows the treatment of intervals of probability, they are not as normatively attractive as the simple form (not to mention that they are technically very complicated). And any way of narrowing the probability ranges will lower the quality of the probability estimate, making it less reliable. This again undermines the normative attractiveness of EUT. So the justificatory gap between ideal and non-ideal EUT mentioned above holds for likelihood estimates as well as utility estimates.

The problem for EUT is not just the justificatory gap. Evidence from psychology and behavioural economics suggests that attempting to use EUT is not always a good strategy in practice. Experimental evidence suggests that agents who violate EUT in certain ways do just as well as, and sometimes better than, those who follow EUT (Arkes, Gigerenzer & Hertwig, 2016; Berg, Eckel & Johnson, 2010). In the field of investment allocation, for example, it has been shown that simply dividing your money evenly between the assets under consideration outperforms sophisticated optimization models in certain conditions (DeMiguel, Garlappi & Uppal, 2007). Generally, using models such as EUT in cases where we do not have precise and good information introduces the possibility of estimation error, while simpler decision procedures are typically biased in some direction (Brighton & Gigerenzer, 2012). In our case, non-ideal EUT risks getting likelihoods and/or utilities wrong, while PP is (or at least may be) biased in the sense that it attaches too much importance to the difference between being over and under the H or K thresholds. In cases such as these, we face a trade-off between minimizing bias and minimizing estimation error. The choice between

these flaws depends on which of the two generates the least bad errors in the concrete decision environment in which we find ourselves.

Steel (2014, Ch. 4) provides evidence that PP's bias is preferable, in the form of a survey of two types of cases: (i) Cases where PP was not applied, even though it was warranted, and (ii) cases where PP was applied, even though it was not warranted. In type-(i) cases, no action was taken despite the existence of evidence that an activity was harmful. In type-(ii) cases, action was taken, but the activity later turned out not to be harmful. Steel argues that the negative consequences of type-(i) cases have been much worse than the negative consequences of type-(ii) cases. The former includes the collapse of fisheries, contamination of sources of ground water, and several hundred thousand deaths, while the latter included the necessity of using other artificial sweeteners than saccharin, and having breast enlargements using other materials than silicone.

It should not be too surprising that PP's bias is less bad than EUT misestimations. As I have argued above, misestimations of irreversible and 'sticky' outcomes may be large, and temporal myopia (as well as explicit time discounting) means that we are prone to underestimate them. On the other hand, excessive precaution is typically limited in time. In cases where the costs of a policy justified by PP are significant, it is highly likely that effort will be put into generating further evidence that shows the relevant activity to be safe, or to designing alternative policies that are less costly. The European Commission demands that precautionary measures be subject to review in light of new evidence, and it requires that an identifiable party be assigned the responsibility for producing more evidence (European Commission, 2000, p. 19-20). So in cases where the original precautionary policy is not warranted, and where it has serious costs, it will (typically) only be in effect for a limited time. Furthermore, the proportionality requirement ensures that precaution does not create high-



er costs than is necessary. The idea that PP asks us to pay a very high price for safety is at least overstated.

## **5. Conclusion**

I have sought to defend PP from the challenge posed by the rationality objections in two ways, a strong-but-shaky and a weak-but-robust argument. §4.1 provides the strong but shaky argument, according to which PP may be (part of) the correct normative theory for risky choices. §4.2 provides the weak-but-robust argument that non-ideal EUT and PP are equally close to ideal EUT – assumed to be the correct normative theory – and that there are reasons to prefer using PP as a decision procedure in these cases. In either case, it is not irrational to use PP, since it is either a way of correctly taking all the reasons that bear on risky choice into account, or a pragmatic decision procedure that realizes our goals in risky choice at least as well, and probably better, than its realistic rivals.

## **References**

- Abdellaoui, M. (2009). Rank-Dependent Utility. In P. Anand, P.K. Pattanaik & C. Puppe (Eds.), *The Handbook of Rational and Social Choice*. Oxford: Oxford University Press.
- Arkes, H., Gigerenzer, G. & Hertwig, G. (2016). How Bad Is Incoherence? *Decision*, 3(1), 20-39.
- Berg, N., Eckel, C. & Johnson, C. (2010). *Inconsistency Pays? Time-inconsistent subjects and EU violators earn more* (MPRA Working Paper). Retrieved from <https://mpa.ub.uni-muenchen.de/26589/>
- Brighton, H. & Gigerenzer, G. (2012). Homo Heuristicus and the Bias-Variance Dilemma. In J. Schulkin (Ed.): *Action, Perception and the Brain: Adaptation and Cephalic Expression*. New York: Palgrave Macmillan.
- Brown, C. (2011). Two Kinds of Holism About Values, *The Philosophical Quarterly*, 57, 456-463.
- DeMiguel, V., Garlappi, L. & Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22(5), 1915-1953.
- European Commission (2000, February 2). *Communication from the Commission on the Precautionary Principle*. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52000DC0001&from=EN>

FOE, ICTA & the ETC Group (2012, March 13). *Principles for the Oversight of Synthetic Biology*.

Retrieved from [http://libcloud.s3.amazonaws.com/93/ae/9/2287/2/Principles\\_for\\_the\\_oversight\\_of\\_synthetic\\_biology.pdf](http://libcloud.s3.amazonaws.com/93/ae/9/2287/2/Principles_for_the_oversight_of_synthetic_biology.pdf)

Goodin, R. (1985). Nuclear Disarmament and Moral Certainty, *Ethics*, 95(3), 641-658.

Harbo, T.-I. (2010). The Function of the Proportionality Principle in EU Law, *European Law Journal*, 16(2), 158-185.

John, S. (2007). How to take deontological concerns seriously in risk–cost–benefit analysis: a re-interpretation of the precautionary principle, *Journal of Medical Ethics*, 33, 221-224.

Lipstich, M. & Ingleby, T.V. (2015). Reply to “Studies on Influenza Virus Transmission Between Ferrets: the Public Health Risks Revisited”, *mBio*, 6(1), 1-6.

Manson, N.A. (2002). Formulating the Precautionary Principle, *Environmental Ethics*, 24, 263-274.

McMahan, J. (1986). Nuclear Deterrence and Future Generations. In A. Cohen & S. Lee (Eds.): *Nuclear Weapons and the Future of Humanity: The Fundamental Questions*. Totowa, NJ: Rowman & Allanheld.

McMahan, J. (2002). *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.

- Pamlin, D. & Armstrong, S. (2015). *12 Risks that threaten human civilization: The case for a new risk category*. Global Challenges Foundation. Retrieved from <http://globalchallenges.org/wp-content/uploads/12-Risks-with-infinite-impact.pdf>
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Peterson, M. (2006). The Precautionary Principle is Incoherent, *Risk Analysis*, 26(3), 595-601.
- Peterson, M. (2008). *Non-Bayesian Decision Theory: Beliefs and Desires as Reasons for Action*. Dordrecht: Springer.
- Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge: Cambridge University Press.
- Randall, A. (2011). *Risk and Precaution*. Cambridge: Cambridge University Press.
- Rozell, D.J. (2015). Assessing and Managing the Risks of Potential Pandemic Pathogen Research, *mBiom* 6(4), 1-4.
- Sandin, P. (1999). Dimensions of the Precautionary Principle, *Human and Ecological Risk Assessment*, 5(5), 889-907.
- Schroeder, M. (2011). Holism, Weight and Undercutting, *Noûs*, 45(2), 328-344.

Shue, H. (2015). Uncertainty as the Reason for Action: Last Opportunity and Future Climate Disaster, *Global Justice*, 8(2), 86-103.

Steel, D. (2013). The Precautionary Principle and the Dilemma Objection, *Ethics, Policy and Environment*, 16(3), 321-340.

Steel, D. (2014). *Philosophy and the Precautionary Principle: Science, Evidence and Environmental Policy*. Cambridge: Cambridge University Press.

Sunstein, C.R. (2005). *Laws of Fear: Beyond the Precautionary Principle*. Cambridge: Cambridge University Press.

Trouwborst, A. (2006). *Precautionary Rights and Duties of States*. Leiden: Martinus Nijhoff.

UNEP (1992). *Rio Declaration on Environment and Development*. Retrieved from <http://www.unep.org/documents.multilingual/default.asp?documentid=78&articleid=11>



## Article 4 On the Cognitive Argument for Cost-Benefit Analysis

### 1. Introduction

For many of us, a moment's self-reflection would reveal that we are far from perfect when it comes to reasoning about probabilities, uncertainties and risk. In recent decades, cognitive psychologists have produced further evidence that confirms this. For example, people tend to neglect the base-rate when judging the likelihood of that a hypothesis is true given some evidence (e.g. the likelihood that a patient has HIV given that an HIV test has come out positive). But base rates matter a lot. Suppose, for example, that the relevant test catches 99% of all HIV cases, and only has a false positive rate of 1% (that is, 1% of the HIV-negative people who are tested show up as positive). Most people would assume that it is overwhelmingly likely that they have HIV, given that the test is positive. But suppose now that the base rate for HIV – the proportion of people in the tested population who as a matter of fact have HIV – is fairly low, say 1/10,000. In that case, the probability of having HIV given that the test is positive is only about 1%.

The tendency to ignore base rates is appropriately known as the base rate fallacy. Another famous mistake in reasoning is the conjunction fallacy, wherein people judge the probability of a conjunction to be higher than the probability of an individual conjunct. In a famous example, subjects were given this short description of a woman: "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations". Subjects were then asked to judge whether it is more probable that Linda is a bank teller or that she is a bank teller and an active feminist. The majority judged that the latter is more probable, which is logically impossible (since in all cases

where Linda is a feminist bank teller, she is also a bank teller *simpliciter*). The base rate fallacy and the conjunction fallacy are perhaps the most famous of the many systematic errors, or cognitive biases, that psychologists have found in ordinary people's reasoning and decision-making, but there are many others.<sup>1</sup>

The aim of most of the work done by psychologists, behavioural economists and others concerning judgment and decision making has had the descriptive aim of understanding how people in fact make judgments and decisions. But more recently, some have begun to draw explicitly normative conclusions from the psychological findings. Probably the most famous example is libertarian paternalism, or 'nudging', which consists in the design of 'choice architecture' – the general set of circumstance in which choices are made – in a way that harnesses biases in a way that produces better choices by the choosers' own lights. For example, default enrolment in pension plans cause people to save more, even though they are free to opt out. Nudging presents an alternative to better-established forms of policy intervention that aims to alter behaviour, such as informational campaigns or economic incentives.

My topic here is not individual behaviour, but society-level regulation of activities that carry a risk of harm to people's health or the environment. In this field, too, psychological findings have been used as the basis for normative arguments. Over the last 15 years, Cass Sunstein (one of the originators of nudging) has argued in a number of writings that there is a cognitive argument for the use of cost-benefit analysis as a basis for social risk regulation.<sup>2</sup> It is safe to say that Sunstein's arguments have been influential in practice, at

---

<sup>1</sup> Wikipedia's "List of Cognitive Biases" includes well over 150 different biases, from "ambiguity aversion" through "the IKEA effect" to the "zero-sum heuristic".

([https://en.wikipedia.org/wiki/List\\_of\\_cognitive\\_biases](https://en.wikipedia.org/wiki/List_of_cognitive_biases), accessed July 14 2016).

<sup>2</sup> Sunstein is not alone in using our cognitive shortcoming as a premise in a case for cost-benefit analysis. For example, Alan Gibbard argues that "something like risk-cost-benefit analysis" is needed in



least in the United States: Sunstein served for three years as Administrator of the Office of Information and Regulatory Affairs (OIRA) during the first term of Barack Obama's presidency. The OIRA is responsible for reviewing regulations from the different agencies of the US government, such as the Environmental Protection Agency (EPA), the Occupational Health and Safety Administration (OSHA) and the Food and Drug Administration (FDA), and by Sunstein's own and others' accounts applied cost-benefit analysis widely under his administration (Heinzerling, 2014; 2015; McGarity, 2013; Sunstein, 2013; 2014). My aim in this paper is to challenge Sunstein's cognitive argument: I will argue that the aim of correcting our cognitive shortcomings in dealing with risk does not warrant the use of cost-benefit analysis as a basis for risk regulation.

## **2. Sunstein's argument**

In a nutshell, Sunstein argues that our judgment and reasoning about cases where risk is involved is biased in various ways, and that cost-benefit analysis can act as a corrective to these biases. Since cost-benefit analysis works as a mechanism for correcting errors, it is justified without reference to any controversial moral views, such as consequentialism or the view that economic efficiency is the sole goal of policy. In Sunstein's words the cognitive argument should lead to an "incompletely theorized agreement" between proponents of most normative outlooks on the use of cost-benefit analysis (Sunstein, 2002, p. 99). In order to understand the cognitive argument, we need to answer three questions: (1) What is cost-benefit analysis? (2) What are the relevant cognitive biases? (3) How does cost-benefit analysis correct these biases?

---

order to "regiment our judgments about risk, and so to avoid the blatant irrationalities of unaided common sense" (Gibbard, 1986: 94).

## 2.1 What is cost-benefit analysis?

At the most general level, cost-benefit analysis (from now on CBA) is a method for determining what the overall societal costs and benefits are of some policy, with a view to judging whether that policy should be carried out, or, where there are several alternatives, *which* policy should be chosen. For risk regulation, CBA would thus consist in an estimate of the value of the harm that regulation prevents and of the costs of implementing regulation.

In order to be able to compare costs and benefits, it is necessary to measure them on a common scale. Typically, that scale is money. That means monetizing the benefits, which may include, among other things, the prevention of premature deaths or other health problems or the preservation of ecosystems. I will follow Sunstein in primarily discussing the risk of premature death. The standard method for generating monetized values for these is by estimating ordinary people's willingness-to-pay (WTP), or in some cases their willingness-to-accept (WTA) for these things. That is, the analyst estimates how much people are willing to pay in order to avoid a certain risk of death or the destruction of an ecosystem (WTP), or alternatively how much they would need to be paid in order to allow themselves to be subjected to a given risk, or to allow an ecosystem to be destroyed (WTA). Since these things are not traded on a market, there are no market prices. Therefore, WTP/WTA must be estimated from relevant market data – most importantly the extra wages workers in risky jobs are paid – or by surveys that directly ask people what their WTP/WTA for some good is.

For analytical purposes it is useful to divide CBA into three aspects:

- (1) Determining what effects a given policy will have – how many lives would be saved, what other health problems would be avoided, which ecosystems would be preserved, how much would it cost industry to implement, how much prices would rise or wages fall etc.

(2) Determining the value of each of these effects using some value metric that allows for direct comparison between disparate effects.

(3) Using money (including WTP/WTB) as the relevant value metric.

I will refer to these as *qualitative analysis*, *commensuration* and *monetization*, respectively.

Commensuration is often described as *quantification* (including by Sunstein). This is misleading, since the description of effects in the 'qualitative' analysis could very easily be quantified, and since the purpose of going beyond qualitative analysis is not simply to put a number on effects, but to find a number that makes effects directly comparable. I will therefore talk of commensuration.

CBA as described here is an analysis of the effects of policy options. While it is tempting to assume that a CBA directly determines what policy should be pursued – i.e. the one with the highest benefits net of costs – this is not implied by CBA itself. And neither does Sunstein advocate this 'hard' use of CBA as a decision rule. Instead, he favours of a 'soft' use of CBA. Most often, he describes CBA as a mere *informational input* into the decision-making process. Such language suggests that CBA is entirely divorced from decision rules – the analysis is just available, providing information to be used or not by decision-makers. But in fact Sunstein does want CBA to be integrated into decision rules. For my purposes, two aspects of Sunstein's concrete policy proposals are important (Sunstein, 2002, pp. 110-113). First, he argues that agencies charged with risk regulation should be required to show that the benefits of a regulation justify the costs. It is somewhat unclear whether Sunstein believes there should be limits to this requirement, and if so what they should be. While he accepts that CBA should not be used where the law explicitly disallows it, he never says anything about whether there are circumstances in which the law *should* disallow CBA. He also believes that the law should as far as possible be interpreted as allowing CBA (Sunstein

2002, Ch. 8).<sup>3</sup> The requirement to show that benefits justify costs is presumptive only; it should be possible for regulative agencies to go ahead with regulations whose costs exceed their benefits, but only “on the basis of a publicly articulated explanation”, e.g. in court (Sunstein, 2002, p. 112). Second, Sunstein argues that agencies should be legally required to use presumptive floors and ceilings for the value of a statistical life (e.g. a floor of \$2 million and a ceiling of \$10 million). These are presumptive in the sense that various “qualitative factors” should be allowed to push valuations above the ceiling or below the floors (I will return to these qualitative factors below). As before, the presumption can only be overridden on the basis of an explicit explanation when the regulation comes under review. The presence of the presumptions and demands for explanation leaves it vague what the ultimate effects of Sunstein’s proposals would be, since no criteria are mentioned for when an explanation for diverging from the presumptions is adequate. This might well be left up to the discretion of judges or civil servants.

## **2.2 Cognitive biases and the corrective effects of cost-benefit analysis**

As mentioned, Sunstein argues that a series of cognitive biases affect our thinking about risk, and that CBA corrects for these. The list of biases differs somewhat between the different writings. Here I attempt to systematize Sunstein’s claims by dividing them into four overarching types of error. These categories are not without overlaps and interconnections, but they seem to me to represent importantly different problems in people’s reasoning.

A minor aside before I present Sunstein’s argument: There is an on-going (and heated) debate within psychology and related sciences about the normative status of the phe-

---

<sup>3</sup> According to an account from a high-level EPA officer at the time, the OIRA under Sunstein’s administration pushed this principle to the limit of legality, and perhaps beyond (Heinzerling, 2014, pp. 25-28).

nomena described below. There is general agreement that human beings exhibit ‘bounded rationality’ – that we use reasoning schemata that are much simpler than those developed by normative theories of practical and theoretical rationality (such as expected utility theory or Bayesian epistemology). Such schemata are known as heuristics. It is also generally agreed that the use of such heuristics lead to biases, i.e. to systematic divergences from the normative theories. But here agreement ends. One school, the heuristics-and-biases school, holds that biases are *errors*, and that using heuristics is only normatively justified when cognitive resources (e.g. time or computing power) is limited. This school includes the Nobel prize winner Daniel Kahneman and his long-time collaborator Amos Tversky, as well as many others – including Sunstein. The other school, known as the adaptive rationality school, views heuristics much more positively. Proponents of this school generally emphasise the importance of matching the decision-making procedure to the environment in which it is to be used. In many decision environments, heuristics yield better results than the standard normative theories of rationality (see Gigerenzer & Selten, 2001). My arguments below do not depend on either of these views of heuristics (and biases) being the correct one, although I do suggest one way in which the rival school would be critical of Sunstein’s proposals.

### **2.2.1 False beliefs about the magnitude of risks**

A risk can be understood as a combination of some *bad outcome* that might occur and the *likelihood* or probability that this outcome does occur. For example, a certain level of exposure to some substance might entail a 1/10,000 chance of developing a fatal cancer. The magnitude of a risk is the product of the bad outcome and the likelihood. As noted, I will focus on cases where the bad outcome in question is premature death. In that case, the magnitude of a risk is equivalent to the likelihood of premature death, and, when summed

over a population, to the annual number of fatalities. Sunstein suggests that several cognitive mechanisms lead people to form systematically wrong beliefs about the magnitude of various risks.

The first such mechanism is the availability heuristic (Sunstein, 2002, pp. 33-35; 2005, pp. 36-39). The availability heuristic is at play when people attempt to estimate the size or frequency of a class or event (such as how many people get divorced after the age of 60). It operates by substituting a judgment of how easily an example of the class or event comes to mind for the judgment of how large the class is or how frequent the event is. The heuristic works in one or both of two ways. Either the frequency is judged by *how many* examples one can think of, or by *how easy* it is to think of an example (Kahneman, 2011, Ch. 12). In a seminal study, the availability heuristic was invoked to explain observed biases in the judgment of risks, framed in terms of “the frequency of lethal events” from various causes (Liechtenstein et al., 1978). The study found (i) that people overestimated low frequency events and underestimated high frequency events, and (ii) that the judged frequency of pairs of events with the same actual frequency differed substantially. The availability heuristic can account for both biases. The overestimation/underestimation bias can be explained by the fact that differences in the number of instances one can recall are much smaller than differences in frequency. The pairwise incongruities was explained by Liechtenstein et al. by the fact that events that were judged as more frequent tended to be highly publicized and emotionally salient ones, such as accidents or cancer, while events with a low judged frequency tended to be “undramatic, quiet killers”, such as diabetes. However, alternative possible explanations for the biases exist, and it is far from settled what drives people’s estimations of frequencies (see Hertwig, Pachur & Kurzenhäuser, 2005).

The second mechanism is really a set of mechanisms of social transmission of information (Sunstein, 2005, Ch. 4). One such mechanism is a cascade. In a cascade, people who

have no strong basis for belief in either direction regarding some proposition – such as that a certain chemical is dangerous – use other peoples' stated beliefs as evidence. Hence A may come to believe that the chemical is dangerous *merely* because B believes so. This effect then expands to others, and is amplified by the fact that a larger number of people now (seem to) believe that the chemical is dangerous. Another mechanism is group polarization, the phenomenon that people tend to move towards more extreme positions as a result of deliberation. In the case of risk, this should lead to the formation of groups that greatly overestimate some specific risk, and to other groups that greatly underestimate other specific risks. A third mechanism is the distortion of information when it is passed on from one person to another, Chinese-whispers-style (see Moussaïd, Brighton & Gaissmaier, 2015). Generally, then, social dynamics may amplify already erroneous beliefs about the magnitude of various risks.

The third cause of systematically false beliefs is emotional responses to risks (Sunstein 2002, pp. 43-47; 2005, pp. 39-41). Sunstein discusses two types of effect that result from emotional processing of information. The first is the affect heuristic. Like the availability heuristic, the affect heuristic is a way of generating a judgment by consulting a related but different phenomenon – in this instance one's general affective response. In the context of risk, this leads people to judge that an activity that they believe to be highly beneficial also carries low risk, and conversely that an activity that they believe to be risky also has low benefits (Alhakami & Slovic, 1994).

The second effect of emotional processing is probability distortion and neglect. According to Sunstein, when we are contemplating a high-affect outcome, we tend to neglect the probability altogether and focus solely on the outcome itself. This accounts both for exaggerated fear of low-probability bad things – of dying in a plane crash, say – and for exaggerated hope of low-probability good things – such as winning the lottery. A review of

studies on emotions' role in decision-making (Volz & Hertwig, 2016) showed that affect increases the non-linear weighing of probabilities in decision (increasing the importance of low-probability events and decreasing the importance of high-probability events). Interestingly, the evidence also suggests that affect triggers the use of simple heuristics, such as a maximin rule (which neglects probability as Sunstein suggests). This finding is consistent with other evidence that a substantial number of people ignore probabilities when making choices involving high-affect outcomes (Suter, Pachur & Hertwig, 2015).

Sunstein cites two further, more specialized mechanisms that lead to false beliefs about risk magnitudes. The first is the general belief that nature is benevolent. As a result, people tend to believe that risks from natural entities are less severe than risks from synthetic or man-made entities (Sunstein, 2005, pp. 44-45). The second is what Sunstein, following Slovic and others (Slovic, 2000, Ch. 18), calls "intuitive toxicology". Intuitive toxicology consists of a set of general beliefs about how toxic agents work, including that a chemical is either toxic or not, regardless of the dose and that it is possible to achieve zero risk exposure from chemicals (Sunstein, 2002, pp. 35-37). Since it is not generally the case that natural substances pose less risk than synthetic ones, or that toxicity is an all-or-nothing matter, these two mechanisms lead to false beliefs about the magnitude of risks caused by exposure to various chemicals in particular.

The general effect of all of the mechanisms described above is that people's beliefs about the magnitude of risks caused by different activities and phenomena are likely to be systematically wrong. CBA corrects this by highlighting the actual magnitude of various risks, as well as by showing precisely how much risk-reduction a given policy would achieve.



### **2.2.2 Neglecting and undervaluing the costs of regulation**

The activities that produce risks typically also have benefits – otherwise human beings would likely not have engaged in them in the first place. According to Sunstein, ordinary people tend to neglect these benefits when they evaluate risky activities. Since the costs and benefits of regulation are partly the mirror image of the costs and benefits of the regulated activity, this means that people tend to neglect some *costs* of regulation. One important reason for believing that people neglect costs is the inverse relationship between perceived benefit and perceived risk described above – i.e. the tendency for people to view activities that are risky as having few benefits (and vice versa). Apart from this general finding, Sunstein puts forward two examples where alerting people to the costs of a policy caused a major shift in attitude towards that policy. One is the removal of asbestos from New York City public schools. Initially, this policy was supported by the parents, but their enthusiasm dropped markedly when they were made aware that schools would need to be closed for several weeks, which was inconvenient to them. The other example is the Kyoto Protocol. According to a survey, 59% of Americans supported the Protocol, with 21% opposed. At the same time, however, 52% said they opposed the Protocol if it would cost the average household an extra \$50 per month (Sunstein 2002, p. 42; 2005, p. 48). In Sunstein's phrase, the costs of regulation are often "off-screen" when people are evaluating policies for dealing with risks.

Sunstein also invokes loss aversion to explain why costs of regulation are neglected or at least undervalued. Loss aversion is the tendency to see a loss from the status quo as more undesirable than a gain from the status quo of the same magnitude is desirable. For example, if offered a bet at 50/50 odds (e.g. a coin toss where you win \$X if the coin comes up heads and lose \$Y if the coin comes up tails), people tend to require the possible winnings to be at least twice as high as the possible loss before they accept the bet (Fox, Erner &

Walters, 2015). Sunstein argues that loss aversion does not merely lead to a lower valuation of equivalent gains and losses, but also gives rise to two biases in people's attitude to risks: (i) People tend to focus mostly on newly introduced risks (and aggravations of existing risks) and focus much less on the benefits that are foregone as a result of regulation; (ii) people are more willing to tolerate familiar risks than new ones (Sunstein, 2005, pp. 42-43).

All in all, Sunstein argues, there is a tendency for people to ignore the costs of regulatory policies (and to ignore the benefits of engaging in a risk-producing activity). The benefits of regulation are "on-screen" while the costs are "off-screen". The asbestos and Kyoto Protocol examples suggest that merely putting costs on-screen can alter people's opinions about policies. CBA corrects cost neglect because all effects of regulation, costs as well as benefits, are present on-screen in a CBA.

### ***2.2.3 Unintended systemic effects and risk-risk trade-offs***

Policy interventions do not occur in a vacuum, but are embedded in a large and complex economic and (where environmental regulation is concerned) natural system. Policies may consequently have unintended effects elsewhere in the system. When people are evaluating different policies, they often ignore such unintended systemic effects and focus solely on the intended effects of the intervention. Sunstein cites a set of experiments using computer simulations to study how people go about trying to solve problems that some people face – e.g. the problems faced by semi-nomadic peoples in West Africa (Dörner, 1990). In these simulations, experimental subjects frequently produced new and worse problems than the ones they were meant to solve. Among the reasons for these failures is the neglect of side- and long-term effects, lack of appreciation for how a system develops over time, and uncoordinated implementation of policies that are mutually counterproductive.

In the case of risk regulation, an especially important phenomenon is the creation of separate substitute risks by the very regulation that is meant to reduce risks, thus generating a “risk-risk trade-off” (sometimes, for health risks, described as a health-health trade-off). Sunstein provides several examples: Fuel economy standards may lead to the production and use of smaller cars that are less safe in crashes; banning asbestos necessitates the use of other, less effective materials in the production of car brakes, rendering the brakes more likely to fail; banning DDT increases malaria risk; and lowering the level of ground-level ozone raises the risk of developing skin cancer and cataracts from sun exposure (Sunstein, 2002, p. 39; 2005, p. 32). He adds to this a general claim, namely that *any* regulation will produce substitute risk via its costs, since lower wealth is generally associated with higher mortality risk (Sunstein, 2002, p. 40; 2005, p. 33).

According to Sunstein, then, people neglect unintended systemic effects. And to the extent that these effects consist in the creation of new risks, the regulation in question may in fact end up raising overall risk rather than reducing it. CBA is a corrective to neglect of systemic effects since it takes a global or system-wide view and includes all effects of regulation in the analysis.

#### **2.2.4 Incoherent valuation of risks**

All of the above phenomena can lead to incoherent valuation of statistical lives – i.e. the assignment of (widely) diverging values to reducing risks that are statistically expected to claim the same number of lives. If we believe we are saving a greater (or smaller) number of lives than we are in fact, the value of each *actual* life saved is likely to be higher (or lower) than it otherwise would be. If we ignore some of the actual cost of saving a given number of lives, costs are likely to be higher. If we are loss averse, we are likely to devote more resources to avoiding ‘new’ deaths rather than to saving lives already projected to be lost.

And if regulation creates substitute risks, we may even find that we have paid to have more people die, rather than less.

Apart from these causes, Sunstein adds two further possible reasons for incoherent valuations. The first is separate evaluation of policies. When we evaluate a policy in isolation, we cannot anchor our judgments in anything outside that policy. Consequently, separate evaluation of two different policies might lead to valuations that seem wrong when we look at both policies together. For example, a study examined people's willingness to pay for two public causes: Cleaning-up of polluted breeding grounds for dolphins, and a skin cancer screening program for workers exposed to the sun for many hours a day. Separate evaluation generated almost identical willingness to pay for the two causes, while joint evaluation generated a much higher willingness to pay for the cancer screening program (Sunstein, Kahneman, Shkade & Ritov, 2002, pp. 1174-1178).

The second is what Sunstein calls "the proportionality effect" (Sunstein, 2002, pp. 47-48). This effect consists in a preference for saving a larger proportion of people in the reference class rather than saving a larger absolute number. Suppose for example that the people living in a village of 100 face a 1-in-20 risk of death, and that the people living in a city of 10,000,000 face a 1-in-1,000,000 risk of death. Many would be deemed it more important to avert the former risk than the latter, although by doing so we would only be saving half the number of lives (five instead of ten). In an early study (Jenni & Loewenstein, 1997), the proportionality effect was found to be the best explanation of the so-called "identifiable victim effect", i.e. the tendency to devote much greater resources to rescuing the lives of identified persons that are immediately at stake (such as trapped miners or a critically ill person) than to elimination of risks (such as increased safety measures in the mine or preventive interventions in health). The upshot is that the amount of resources devoted per statistical life

saved is likely to be higher for small populations facing somewhat higher probability of death than for larger populations facing somewhat lower probability of death.

According to Sunstein, CBA provides a corrective for incoherent evaluations in two ways. First, it provides information about how many lives are saved and at what cost. Second, it forces policy evaluation that refers to other policies as well, and thus prevents isolated evaluation. Apart from this, the correction that CBA provides for all the other cognitive biases presumably also contribute to the eradication of incoherence.

### **3. Is cost-benefit analysis necessary for correcting the biases?**

Recall that cost benefit analysis can be disaggregated into three parts, qualitative analysis, commensuration and monetization. Recall also that Sunstein advocates two concrete institutional implementations of CBA, namely a legal requirement that regulatory agencies must show that the benefits of a regulation justifies its costs, and the use of legally mandated presumptive floors and ceilings for the value of a statistical life (from now on I will abbreviate this VSL). Which aspects of CBA are necessary for correcting the various cognitive biases described above? Or in other words, which aspects of CBA are warranted by the goal of disciplining risk regulation so as to avoid systematic errors?

Before answering these questions, it might be instructive to consider which aspects of CBA Sunstein's two institutional suggestions rely on, or rather what aspects of CBA would be *de facto* implemented if the institutional suggestions were to be implemented. Take first the requirement that benefits must be shown to justify costs. In order to make any such judgement, it is necessary to commensurate the benefits and costs at least to a minimal degree. Any judgment that benefits *do* (or do not) exceed costs in a given instance can be read backwards to imply some set of relative values. For example, the judgment that the benefits of a regulation that saves 10 statistical lives (and has no other benefits) at the cost of \$200

million to taxpayers is justified implies that, at least in this instance, the 10 statistical lives are worth more than the \$200 million – we are behaving *as if* we take the VSL to be at least \$20 million.

This minimal or as-if commensuration falls short of the commensuration envisaged by CBA in two ways. First, it does not pass judgment on the relative values of the different *constituent* costs and benefits – it merely requires a judgment that *all* the benefits, taken as a whole, are larger than *all* the costs, taken as a whole. Second, the relative values implied by a judgment in one case can in principle be wholly independent of the relative values that are ‘used’ in another case.<sup>4</sup> It is fairly clear that what Sunstein has in mind is that the benefits *as defined by CBA* must justify the costs *as defined by CBA* (with the caveat that this is only a presumptive demand, of undefined strength, for ultimate justification of a regulation). So commensuration of a stronger kind – using valuations of constituent costs and benefits that are the same across cases – would be implemented through Sunstein’s first institutional suggestions.

Presumptive floors and ceilings for the VSL precisely establish such a set of relative values of constituent costs and benefits that are determined prior to the specific cases, and that are supposed to hold generally (again with the caveat that the presumption can, at least in principle, be overridden by ‘qualitative factors’). In principle, another metric than money could have been used – Sunstein describes monetization as a mere “pragmatic tool” (Sunstein, 2002, p. 111). But in practice, monetary equivalents are the only commensuration tool that is ever seriously considered in CBA, by Sunstein as well as by any other proponent. Furthermore, the specific floors and ceilings that Sunstein suggests are derived from

---

<sup>4</sup> It might plausibly be argued that this second feature means that minimal commensuration is not useful for *justifying* regulations. At the very least, justifications should be constrained by the precedents set by judgments in similar cases.

WTP/WTB (Sunstein, 2005, Ch. 6). So Sunstein's two policy proposals imply the use of commensuration in the sense of an establishment of relative values for the constituent benefits and costs of regulation, and in practice uses monetization and some form of WTP/WTB to come up with a set of such relative values. It goes beyond mere qualitative analysis of what costs and benefits a policy will have.

Return now to the question of whether the cognitive biases described in the previous section give reason to implement CBA, in particular the Sunstein's two institutional suggestions. Consider first false beliefs about the magnitude of risks. To avoid false beliefs, we need only an analysis that shows us what the actual magnitudes of risks are. In other words, we only need qualitative analysis. Similarly for the neglect of costs and the neglect of substitute risks: It is not necessary to commensurate these with the benefits of regulation in order to bring them onto our 'viewscreen'. So the desire to avoid false beliefs and to ensure consideration of all effects of regulation does not in itself justify the use of CBA as Sunstein wants it used. (For example, demand that we use floors and ceilings for the VSL does nothing to correct the underestimation of fatalities from undramatic, quiet killers such as diabetes). A large part of Sunstein's cognitive argument thus trades on a false dilemma between CBA and what Sunstein himself describes as "uninformed stabs in the dark" (Sunstein, 2002, p. 6). It is only if all alternatives to CBA neglect or refuse to even consider relevant information about the magnitude of risks and the effects of regulation that CBA is supported. But this is implausible. Sunstein implicitly compares CBA only with relying more or less directly on the possibly uninformed views of the public, which was never a serious candidate for a policy of risk regulation.

There is a possible reading of Sunstein's claims that does put him into conflict with more than uninformed stabs in the dark. On this reading, Sunstein argues that all the possible effects that may be neglected – regulatory costs, substitute risks and systemic effects –

*must* be considered in each instance of regulation. This goes beyond the demand that we do not use wrong information by demanding that we base our decision on correct information about all effects. This seems plausible enough initially. But it ignores the fact that information is not simply available to be looked at – it is the product of active investigation of the world by scientific or other means. A requirement to look at the facts may therefore easily slide into a requirement to *produce* facts. This is where the rival school within the psychology of heuristics have a point against Sunstein. Drawing especially on the ideas of Herbert Simon, proponents of this school argue that producing further information is often not worth the effort, since the value it adds to the quality of the decision is limited. In the case of risk regulation, demands that certain types of effects be considered can lead to delays in implementing regulations (for example, the demand that substitute cancer risks be considered delayed EPA regulation of ground-level ozone for several years). It may therefore be reasonable to ignore some (kinds of) information when deciding whether and how to regulate some risk. *When* it is reasonable and *what* kinds of information to ignore are complicated questions that I will not go into here.

#### **4. Incoherence and cognitive error**

The cognitive errors that manifest themselves in false beliefs and neglect of relevant factors (especially the costs of regulation) are in principle avoided simply through the use of qualitative analysis. Consequently Sunstein's two institutional proposals are not warranted by the desire to avoid these errors. With respect to coherence, however, the two proposals seem *prima facie* warranted. Requiring that regulations can be shown to generate higher benefits than costs and using floors and ceilings for the VSL directly generates more coherence, since it limits the possible variations in costs per statistical life saved that are possible. However, incoherence is less obviously irrational than basing regulation on false beliefs or



neglecting relevant information. There are three possible sets of reasons for thinking that coherence should be an uncontroversial goal of policy: Methodological, instrumental, and moral reasons. In this section I will examine each of these sets of reasons and argue that the coherence that is warranted by each set of reasons is not the coherence that CBA offers.

#### **4.1 Methodological reasons**

Within moral epistemology, the view that coherence is necessary for our normative view to be justified is commonplace. The overarching idea is that we should compare our judgments in one case with our judgments in other cases, as well as with more general moral principles. Coherentist methods include the case-based and analogical forms of reasoning that are very common in applied ethics and method of reflective equilibrium that is arguably the dominant view of justification within ethics generally. Sunstein's discussion of separate versus joint evaluation illustrates the virtues of a coherentist view of justification. Joint evaluation clearly increases coherence in the sense of 'coherence' used by moral epistemologists, while separate evaluation does not. As Sunstein shows, separate evaluation can lead to judgments that are reversed after joint evaluation, and that are therefore not justified on a coherentist view of justification. However, coherentist justification is not the same as what I have described as coherence above. There, coherence denotes the use of the same VSL (or rather, the use of floors and ceilings) across all regulations. We cannot assume that coherence in the moral-epistemological sense will result in the view that coherence in the CBA sense is the right judgment.

But perhaps the practice of CBA itself constitutes a coherentists method, i.e. is a type of joint evaluation of the VSL. *Prima facie*, it might seem to, since CBA tries to generate a VSL that holds across all policy domains. But looking at the standard method used, it is clear that no joint evaluation is happening. Instead, the VSL is typically extracted from a

single (type of) case, namely the wage premiums workers receive for taking jobs with mortality risks. The comparison between wage premium cases and regulation cases is entirely one way – no argument is ever made that wage premiums are too high, or too low, based on ‘feedback’ from the regulation case. And the VSL itself is not adjusted based on reasoning about what level of expenditure is reasonable in different regulation cases. So the fact that a given VSL is used in CBA gives us no evidence that joint evaluation would produce that VSL. It only gives us evidence that extrapolating from a single separate evaluation gives us that VSL. This is no coherentist method.

A different attempt to justify coherence (in the sense of a set VSL) on methodological grounds is to argue that the valuation of lives would otherwise likely be random or arbitrary (Sunstein, 2005, p. 149). But the mere fact that the implied valuations of life differ widely between regulations does not imply that they are arbitrary; they may instead be the result of a regulatory rationale that does not use anything like a VSL as input. Consider, in particular, a relatively popular alternative approach to regulation, namely technology-based regulation (see McGarity, 2002b, p. 2343-2344). In technology-based regulation, regulators demand that risk is reduced to the maximum degree that is technologically feasible (which normally amounts to less than a complete elimination of the risk), regardless of cost.<sup>5</sup> The implied VSL of a technology-based regulation is determined by the cost of the best available technology and how many lives that technology will save. There is no particular reason to think that the implied VSLs of a set of technology-based regulations will be the same, or will fall within some range. But neither are the VSLs arbitrary – they are the product of a reasoned approach to risk management.

---

<sup>5</sup> Except to the extent that cost is allowed to play a role in determining what is and is not feasible, which it likely will.

## 4.2 Instrumental reasons

Assume that all we care about really is saving as many lives as possible. Then it seems that any way of going about our business that results in fewer lives than possible being saved than is instrumentally irrational. Sunstein argues that current (or perhaps rather pre-CBA era) regulatory policy suffers from exactly that instrumental irrationality. He reproduces a table (which, according to him, “has come to define many discussions of these problems”) showing that the cost per premature death averted for a large set of regulations varied from \$0.1 million to \$92 billion<sup>6</sup> (Sunstein, 2002, p. 30). The conclusion he draws is that the resources the United States government spend to reduce risks are poorly allocated. He cites a study that suggests that reallocating resources could save an additional 60,000 lives annually, holding cost constant, or save \$31 billion, holding lives saved constant (Sunstein, 2002, p. 25; the study cited is Tengs & Graham, 1996). These numbers are somewhat old, not to mention highly controversial (for critiques see Heinzerling, 1998; Parker, 2003). More recent are the net benefits from regulation that Sunstein reports was achieved during his three years as head of the OIRA. According to calculations by the White House Office of Management and Budget (of which OIRA is a part) the early Obama administration managed, through the use of CBA, to increase net benefits from \$3.4 billion annually (under the Bush administration) to \$91.3 billion (Office of Management and Budget, 2012, p. 59; Sunstein, 2013, pp. 33-35).

The first question we must ask is whether it is really true that we could save many more lives by introducing coherence, and that regulatory net benefits have risen dramatically after the use of CBA was increased. With respect to the latter, there may be a measure of question begging involved, since the calculations made by the Office of Management and

---

<sup>6</sup> The \$92 billion figure is a big outlier; the 2nd most expensive regulation spent around \$4.2 billion and the 3rd most expensive ‘only’ \$653 million.

Budgets are based on monetized costs and benefits as recommended in CBA. It is not surprising that a policy aimed at maximizing CBA-defined net benefits in fact increases CBA-defined net benefits. With respect to the former, the mere fact that some regulations spend smaller sums per statistical life saved than others is not sufficient to show that more lives could be saved. First of all, what matters is not actual expenditure per life saved, but the marginal cost of saving an *extra* life for each regulation. There is no general reason to think that actual expenditure is a good indicator of the marginal cost of an extra life saved. For one thing some regulations may have succeeded in eliminating mortality risk altogether, leaving no extra lives available to be saved. For another, the potential of one relatively cheap method of reducing risk may have been reached, and any alternative method may be much more expensive. This would be the case, for example, when there is no further way of reducing the risk from some activity besides banning the activity altogether.<sup>7</sup> Second of all, our ability to save more lives requires that the savings on a relatively expensive regulation can be reallocated to a cheaper regulation. But in many cases, the savings associated with not regulating (or repealing a regulation) accrue to various private actors, be they businesses, workers or consumers. Such savings are not immediately available to 'spend' on other regulations (although they could perhaps in principle be made available).

Suppose we could, in fact, save more lives at the same cost (or save the same number of lives at less cost) by reallocating our regulatory efforts. This fact alone still does not warrant the use of a full CBA that commensurates and monetizes the value of lives saved. Sup-

---

<sup>7</sup> The study by Tengs and Graham cited by Sunstein does take the marginal cost issue into account by using numbers for the "level of implementation" of various interventions (so we can assess whether an intervention has residual life-saving potential). However, those numbers are estimates by anonymous reviewers, and there is no publically available information about what their basis is. Furthermore, more than two thirds of the potential extra lives saved could be found in just two medical interventions, namely providing round-the-clock oxygen for sufferers of hypoxemic obstructive lung disease and influenza vaccines for everyone (Parker 2003: 1362-1363, 1377-1381).

pose we were to hold current expenditure fixed and then go about saving lives in the most cost-efficient way. In that case, the VSL would be determined by what the cost of the final life we could afford to save is. There would be no need to determine the VSL by any external method, such as WTP/MTA. The effect of using an externally determined VSL is to implicitly set our total risk reduction 'budget' – which would be the total expenditure we would incur by saving all lives that could be saved at less than \$X each. There is an allocation-based rationale for using the externally determined VSL, but it has nothing to do with how many lives we could save at a given cost. Rather it has to do with what other goods we can get instead of a statistical life saved. It may be instrumentally irrational to save fewer lives than we could have done for the same amount of money, but it is not necessarily instrumentally irrational to spend any particular amount of money on saving lives. Only if the externally determined VSL captures the *correct* value ratio between risk reduction and everything else will it be instrumentally irrational to spend more than the VSL on saving a statistical life (since we are getting less of overall good than we could have gotten via an alternative allocation of resources).

#### **4.3 Moral reasons**

Moral reasons for coherence would show that using a set VSL is required by morality. Since Sunstein wants to achieve an incompletely theorized agreement on the use of CBA across different moral theories, the moral reasons for coherence must be generally accepted as valid. There are ways of making it sound highly intuitive that the VSL should be the same in all cases. Surely all lives are worth the same; surely it would be unjust to be willing to spend only \$1 million to save a life in one domain and \$20 million to save a life in another. The underlying assumption seems to be that the VSL is a measure of how important the life potentially saved is. If we are using as VSL of \$1 million in the domain of workplace acci-

dents and a VSL of \$20 million in the domain of air traffic, then we are implicitly saying that the life of a worker is only worth one twentieth of the life of an air traveller. However, as I have argued above, the VSL need not be a direct input into decision-making, but can instead be viewed as an unintended consequence of some other rationale for decision (such as technology-based regulation). Below, I offer four types of rationale that are part of plausible moral theories, that do not use VSL as a decision input, and that leads to predictable incoherence in the VSL.

#### **4.3.1 *Ex ante and ex post***

The aim of coherence in the VSL assumes that the morally important phenomenon is saving statistical lives (or avoiding statistical deaths). This is the *ex post* perspective in the ethics of risk. The *ex post* perspective evaluates risk-involving social decisions on the basis of the distribution of burdens and benefits that will arise after the uncertainty has been resolved (e.g. because a chance process has taken its course). But there is an alternative view, namely that it is the *risk* that people face *ex ante* that should be the focus of our attention. The *ex ante* perspective evaluates risky social choices on the basis of the distributions of benefits and burdens *before* the uncertainty has been resolved. For those moral theories that place importance on the distribution of burdens and benefits, the *ex ante* view implies that it is morally worse if 1,000 people bear a 1-in-1,000 risk of death than if 1,000,000 people bear a 1-in-100,000 risk of death, even though the number of statistical lives lost is the same in both cases (see e.g. Frick, 2015; Lenman, 2000).

The *ex ante* view would (for distribution-sensitive theories) therefore lead to the same incoherences in the VSL that Sunstein attributes to the proportionality effect (i.e. the tendency to want to save the larger proportion of the reference groups to which different potential victims belong). So if the *ex ante* view is correct, we should perhaps welcome at least

this form of incoherence. Sunstein seems to acknowledge this fact; the only cognitive *error* that Sunstein finds with regard to the proportionality effect is that it is “operative as an automatic, unreflective intuition” (Sunstein, 2002, p. 48). But avoiding such unreflective use is arguably solved by qualitative analysis alone, at least insofar as such an analysis includes a tally of statistically expected lives saved as well as the magnitude of risk imposed on each risk bearer. The further aspects of CBA, and in particular the use of a set VSL range, do nothing to answer the question of whether, why, and to what extent there should be a non-proportional difference in the value of imposing a 1/10 risk of death per year and a 1/100,000 risk of death per year.

The debate between *ex ante* and *ex post* views is far from settled within the ethics of risk (see Cohen, Daniel & Eyal, 2015). No matter how that debate is eventually settled (if it is), proponents of CBA face a troubling dilemma. A common criticism of CBA is that it puts a price on a human life, and thus implies that lives may be ‘sold’ for a certain dollar amount. The idea that (for example) an industry can legitimately kill 5 people as long as it generates savings to consumers and extra profits of at least \$45 million<sup>8</sup> seems plainly wrong. In order to answer this objection, Sunstein and other CBA proponents often stress that using a VSL does not place a dollar value on a life *per se* but rather on a mortality risk (Sunstein, 2005, p. 137). This suggests that CBA is an *ex ante* view, and hence should not be especially concerned with the simple number of statistical lives saved, nor with using a set VSL (but rather a set value for a 1/10,000 mortality risk, and another set value for a 1/100,000 risk).

So it seems that CBA proponents can either hold on to the demand for coherence in the VSL and accept that *lives*, rather than risks, have a dollar value, or give up the focus on coherence in the VSL and hold on to the idea that it is mortality risks that have a dollar val-

---

<sup>8</sup> Using the VSL of \$9 million currently used by several US government agencies.

ue. Arguing that the VSL coherence is important *and* that mortality risks are priced is not an available option – at least not as long as we want a distributionally sensitive theory. In my view, the more attractive alternative is to give up coherence as an ideal. At any rate, it is not at all obvious that lack of coherence represents a moral failing.

#### **4.3.2 Distribution of costs and benefits**

Different regulations exhibit different distributions of costs and benefits. There are two very simple kinds of case (and a range of unsimple ones in between). The first simple kind of case is one where the beneficiaries of regulation also bear the entire cost. The second simple kind of case is the opposite, where the beneficiaries and the cost bearers are wholly distinct groups. In between are combinations of the two, where different persons incur some of the cost and/or some of the benefit of a regulation. Let cases where the cost is entirely borne by the beneficiaries be called *type A* cases, and cases where beneficiaries and cost bearers are entirely distinct *type B* cases. The standard rationale given by proponents of CBA, including Sunstein, for declining to regulate in cases of type A where the cost exceeds each person's WTP is that it would amount to a "forced exchange" that the so-called beneficiaries would rather have avoided (Sunstein, 2005, pp. 150-153). This rationale does not require the *regulator* to take any stance on how much people's lives are worth – the VSL is an artefact of a rationale – respecting people's right to choose for themselves – that does not consider VSL.

In cases of type B, the forced-exchange rationale is not applicable, least not directly. The beneficiaries are here forced to buy risk reduction, but rather given a 'gift' of reduced risk. If the costs of this gift exceed the beneficiaries' combined WTP, all that means is that there is *another* gift that the beneficiaries would rather have had (namely an amount of money that strictly exceeds their WTP to remove the relevant risk). This *indirect* forced-exchange rationale can only be used to block regulation if one at the same time supports a



transfer of money to would-be beneficiaries of the regulation. But in most cases, that proposal is not on the table – only regulating or not regulating is. Denying regulation on the basis of the indirect forced-exchange rationale would be analogous to not giving your wife flowers because she would rather have chocolate – but then not giving her chocolate either.

So the rationale for using WTP in a type B case must be something else than the forced-exchange rationale used in type A cases. Consequently, using a different VSL than the one implied by WTP in type B cases does not amount to treating the lives saved in type-B cases differently from the lives saved in type-A cases. The different rationales mean that the VSLs implied express qualitatively different judgments. In type A cases, the judgment is mainly about respecting individuals' right to choose and their own judgments of the relative value of risk reduction and money. In type B cases, the judgment may be about how much money we, as a society, should be willing to spend to reduce risks (if costs are borne by taxpayers) or it may be about what burdens we may legitimately place on *others* in order to reduce risks (if costs are borne by other private actors, e.g. businesses or consumers). There is no reason to think that all of these judgments must yield the same number, and hence no reason to expect VSL coherence across the cases. More problematically (for CBA), it seems that there is no reason for using the WTP number, which reflects a respect-for-individuals judgment, in type B cases, which do not concern respect for individuals. Consequently we are left without any basis for setting the VSL at any particular level in these types of case.

#### **4.3.3 Qualitative factors**

One kind of rationale straightforwardly leads to VSL incoherence, namely a rationale according to which it is not true that every saving of a statistical life has the same moral importance. Sunstein acknowledges that there are differences between cases that warrant us-

ing different VSLs. In particular, he argues that deaths that are especially dreaded, such as cancer deaths, should be assigned a higher VSL; that involuntarily borne risks should be assigned a higher VSL than voluntarily borne ones; that large-scale catastrophes may be worse than the mere number of lives lost indicate; and that risks to especially vulnerable or underprivileged groups, such as children or the poor, may be given higher weight. These are the 'qualitative factors' mentioned above. Taking the qualitative factors into account does not amount to treating different people's lives as having different moral worth.

The qualitative factors hide a number of rationales for risk regulation that goes beyond merely saving lives. The idea that we should care more about dreaded deaths reflect the fact that there is more to dying than death itself, e.g. that many ways of dying involve large amounts of suffering. Adjusting the VSL is an adequate way of taking this into account, since the rationale is basically consequentialist – it suggests that one consequence, e.g. death from cancer, is worse than another, e.g. death from a workplace accident. The rationales behind other factors are less obviously honoured by adjusting the VSL. The idea that voluntariness matters reflects the importance of a number of phenomena, including freedom, consent, compensation and personal responsibility. The proper moral role of these factors is complex and disputed (for brief discussions of some of the complications see Thomson, 1986, pp. 169-172 & 188-191). But whatever the role of these factors is precisely, the value of a statistical life is not likely to be part of the description of what difference it makes whether someone consents to bearing a risk or not, is compensated for the risk or not, or bears personal responsibility for being subjected to the risk or not. The cost of avoiding the risk may not even be part of that description at all. Thus we should not even expect coherence-with-adjustments for the VSL if these qualitative factors play the role they should in setting policy on risks.

It is also worth noting that qualitative factors do not only matter for the evaluation of the risk reduction – i.e. the benefits of a regulation – but also for the costs. Once the relevance of factors beyond purely descriptive consequences has been admitted, it seems much less obvious that various systemic effects and substitute risks count against regulations. For example, the increased risk of skin cancer and cataracts from sun exposure that is caused by remove ground-level ozone pollution is easily avoidable by wearing sunscreen, sunglasses or suitable clothes. It thus seems to be voluntarily borne, and it is therefore questionable if it should matter (much) in policy-making. The relevance of the general association between regulatory costs and increased mortality will similarly depend on the details of how costs cause increased mortality, and on what other things (besides not regulating) can be done to avoid it. Recognizing the relevance of qualitative factors thus blurs the line between *neglecting* certain costs of regulation and *rightly treating them as morally irrelevant*.

#### **4.3.4 Counting costs**

Suppose we accept the idea that each saving of a statistical life should be given the same moral importance. The ideal of a coherent VSL does not follow from this principle. The reason is that assigning equal moral importance to each saving of a statistical life is interpreted in CBA as letting every statistical life justify the same regulatory costs. In other words, the VSL measures moral importance in terms of the costs of regulation. If it is not the case that every regulatory cost has the same moral importance, CBAs ideal of coherence does not coincide with assigning equal moral importance to all lives. And it is highly plausible that not all costs are equally morally important. Most basically, the imposition of a cost may in some cases amount to a seizure of illegitimate gains. This is especially true when those who suffer the costs of regulation are also causally responsible for the existence of the risk and

do not (mainly) bear the risk themselves – that is, when they *impose* the risk on others for their own benefit. In many such cases, the costs arguably should not matter at all.

But suppose both the beneficiaries of regulation and those who bear the cost have an equal *prima facie* claim to be helped. Even in such cases, plausible moral theories hold that we should make distinctions between different costs. Many non-consequentialists argue that aggregation is not permitted in cases where the members of one group stand to get a minor benefit while the members of the other group stand to get a major benefit. For example, if we can either save one person from severe pain or save one billion people from suffering a inconvenience of not being able to watch the World Cup final, we should save the one, even if the aggregated benefit to the one billion would be larger (Scanlon, 1998, pp. 235-236). This has implications with respect to the regulation of risk. Suppose group A will bear the costs of a regulation while group B will benefit (by being relieved of bearing a risk). If the cost to each member of A is trivial compared with the risk that the members of B will otherwise bear, then on a non-consequentialist view we should help B – that is, implement the regulation – even if the aggregate benefits to group A are larger (see Lenman, 2000; James, 2012; Frick, 2015). There are many open questions here, such as how *ex ante* and *ex post* claims or burdens should be incorporated. But no matter how these are resolved, contractualism denies that all costs should be taken into account (in the sense that they are made to carry weight in favour of one policy or another). For example, minor price increases that hit a large number of A-people should not be counted when the B-people would face a sizeable risk of death. If risks are regulated in a way consistent with the non-consequentialist view, then we should expect different life savings to have different *aggregated* costs, i.e. different VSLs.

There is a more general problem with the way costs are treated in CBA than that plausible moral theories disagree with it. Those who argue for and practice CBA are much

less attentive to differences on the cost side than to differences on the benefits side. Consider, for example, Sunstein's description of how 36 ideal-type cases of regulation would be (and presumably were) scrutinized using CBA during the process of regulatory review that he administrated (Sunstein, 2014, Ch. 2). The number of questions posed to the cost side is precisely zero – in every case the cost is merely stated as a total aggregate dollar amount. At least one reason why CBA's critics are often especially averse to the monetization aspect of CBA is exactly that it makes the analysis blind to morally important differences on the cost side, and cannot account for morally relevant differences in the trade-offs made in different instances of setting a VSL at \$X. Proponents of CBA (including Sunstein) frequently argue that monetization is merely one pragmatic way of commensurating values. I suppose it could be, but not as long as values that are 'naturally' in a money format are not subjected to any rational scrutiny. If CBA is to be justified on cognitive grounds, the fact that it fails to rationally scrutinize one side of the cost-benefit equation seems to be especially troubling.

## **5. Conclusion**

I have argued above that CBA, and in particular Sunstein's two institutional proposals, goes beyond what is necessary for correcting cognitive biases. The commensuration and monetization aspects of CBA are not necessary for getting rid of false beliefs about the magnitude of risks, or for making sure that all costs of regulation are on-screen; and the coherence for which commensuration and monetization is not justified either on methodological, instrumental or moral grounds. The cognitive argument for CBA thus does not justify CBA beyond qualitative analysis, which is a feature that more or less any decision procedure can incorporate.

## **References**

- Alhakami, A.S. & Slovic, P. (1994). A Psychological Study of the Inverse Relationship Between Perceived Risk and Perceived Benefit, *Risk Analysis*, 14(6), 1085-1096.
- Cohen, I.G., Daniels, N. & Eyal, N. (Eds.) (2015): *Identified Versus Statistical Lives: An Interdisciplinary Perspective*. Oxford: Oxford University Press.
- Dörner, D. (1990). The Logic of Failure, *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 327, 463-473.
- Finkel, A.M. (2014). The Cost of Nothing Trumps the Value of Everything: The Failure of Regulatory Economics to Keep Pace with Improvements in Quantitative Risk Analysis, *Michigan Journal of Environmental and Administrative Law*, 4(1), 91-156.
- Fox, C.R., Erner, C. & Walters, D.J. (2015). Decision Under Risk: From the Field to the Laboratory and Back, in G. Keren & G. Wu (Eds.): *The Wiley Blackwell Handbook of Judgment and Decision Making*. Chichester: John Wiley & Sons.
- Frick, J. (2015). Contractualism and Social Risk, *Philosophy and Public Affairs*, 43(3), 175-223.
- Gibbard, A. (1986). Risk and Value. in D. MacLean (Ed.), *Values at Risk*. Savage, MD: Rowman & Littlefield.
- Gigerenzer, G. & Selten, R. (Eds.) (2001). *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.

Heinzerling, L. (1998). Regulatory Costs of Mythic Proportions, *The Yale Law Journal*, 107(7), 1981-2070.

Heinzerling, L. (2014). Inside EPA: A Former Insider's Reflections on the Relationship Between the Obama EPA and the Obama White House, *Pace Environmental Law Review* 31(1) 325-369.

Heinzerling, L. (2015). Statistical Lives in Environmental Law. In I.G. Cohen, N. Daniels & N. Eyal (Eds.), *Identified Versus Statistical Lives: An Interdisciplinary Perspective*. Oxford: Oxford University Press.

Hertwig, R., Pachur, T. & Kurzenhäuser, S. (2005). Judgments of Risk Frequencies: Tests of Possible Cognitive Mechanisms, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(4), 621-642.

James, A. (2012). Contractualism's (Not So) Slippery Slope, *Legal Theory*, 18(3), 263-292.

Jenni, K.E. & Lowenstein, G. (1997). Explaining the "Identifiable Victim Effect", *Journal of Risk and Uncertainty*, 14, 235-257.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Penguin Books.

Kamm, F.M. (2007). *Intricate Ethics: Rights, Responsibilities and Permissible Harm*. Oxford: Oxford University Press.

Lenman, J. (2000). Contractualism and Risk Imposition, *Politics, Philosophy and Economics*, 7(1), 99-122.

Liechtenstein, S. Slovic, P., Fischhoff, B., Layman, M. & Combs, B. (1978). Judged Frequency of Lethal Events, *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 551-578.

McGarity, T.O. (2002). Professor Sunstein's Fuzzy Math, *The Georgetown Law Journal*, 90, 2341-2377.

McGarity, T.O. (2013). EPA at Helm's Deep: Surviving the Fourth Attack on Environmental Law, *Fordham Environmental Law Review*, 24(2), 205-241.

Moussaïd, M., Brighton, H. & Gaissmaier, W. (2015). The Amplification of Risk in Experimental Diffusion Chains, *PNAS*, 112(18), 5631-5636.

Office of Management and Budget (2012). *2012 Report to Congress on the Benefits and Costs of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities*. Retrieved from [https://www.whitehouse.gov/sites/default/files/omb/inforeg/2012\\_cb/2012\\_cost\\_benefit\\_report.pdf](https://www.whitehouse.gov/sites/default/files/omb/inforeg/2012_cb/2012_cost_benefit_report.pdf)

Parker, R.W. (2003). Grading the Government, *The University of Chicago Law Review*, 70(4), 1345-1486.



Slovic, P. (2000). *The Perception of Risk*. London: Earthscan.

Sunstein, C.R. (2001). Cognition and Cost-Benefit Analysis, *The Journal of Legal Studies*, 29(52), 1059-1103.

Sunstein, C.R. (2002). *Risk and Reason: Safety, Law and the Environment*. Cambridge: Cambridge University Press.

Sunstein, C.R. (2005). *Laws of Fear: Beyond the Precautionary Principle*. Cambridge: Cambridge University Press.

Sunstein, C.R. (2013). *Simpler: The Future of Government*. New York: Simon & Schuster.

Sunstein, C.R. (2014). *Valuing Life: Humanizing the Regulatory State*. Chicago: The University of Chicago Press.

Sunstein, C.R., Kahneman, D., Schkade, D. & Ritov, I. (2002). Predictably Incoherent Judgments, *Stanford Law Review*, 54(7), 1153-1215.

Suter, R.S., Pachur, T. & Hertwig, R. (2015). How Affect Shapes Risky Choice: Distorted Probability Weighing Versus Probability Neglect, *Journal of Behavioral Decision Making*, doi: 10.1002/bdm.1888

Tengs, T.O. & Graham, J.D. (1996). The Opportunity Cost of Haphazard Social Investments in Life-Saving. In R.W. Hahn (Ed.), *Risks, Costs and Lives Saved: Getting Better Results from Regulation*. Oxford: Oxford University Press.

Thomson, J.J. (1986). *Rights, Restitution and Risk: Essays in Moral Theory*. Cambridge, MA: Harvard University Press.

## Resume

Afhandlingen diskuterer de etiske problemstillinger som rejses af syntesebiologi – en ny bioteknologi, hvor forskerne kan designe og bygge organismer der har gavnlige egenskaber.

I indledningen giver jeg et kort overblik over forskningen indenfor syntesebiologien og de samfundsmæssige reaktioner der har været på den. Fokus er på de dele af teknologien der er mest interessant fra et etisk perspektiv, og som har været mest diskuteret i akademiske, politiske og samfundsmæssige kredse. Indledningen giver en introduktion til de fire artikler som udgør afhandlingens hoveddel. Den består desuden af en mere dybdegående diskussion af en væsentlig klasse af indvendinger mod syntesebiologi, nemlig indvendinger der baseres på det menneskelige forhold til andre levende væsener og naturen

I Artikel 1 og 2 diskuterer jeg en række problemer i den måde etiske problemstillinger i syntesebiologi (og nye teknologier generelt) vinkles på. I artikel 1 argumenterer jeg for at det at skabe kunstigt liv er en moralsk væsentlig handling. Jeg kritiserer den forståelse af 'kunstigt liv' og 'moralisk væsentlighed' som anvendes af Douglas, Powell og Savulescu, for at være for snævre og for at lede til blindhed overfor visse indvendinger. I artikel 2 kritiserer jeg et udbredt forsvar for syntesebiologi, baseret på relevante ligheder mellem syntesebiologi og etablerede teknologier, og viser at visse indvendinger kan undvige dette lighedsargument.

I artikel 3 og 4 diskuterer jeg måder at håndtere usikkerhed og risiko på. Artikel 3 forsvarede forsigtighedsprincippet fra kritikere der hævder, at princippet er irrationelt. Artikel 4 kritiserer Sunstein's argument for cost-benefit analyse som en korrektion af kognitive fejl. Jeg viser at kun ikke-centrale dele af cost-benefit analysen kan forsvares på denne måde

## Summary

The dissertation explores ethical issues concerning synthetic biology, an emerging biotechnology whereby researchers design and construct organisms that are useful for human beings.

The introduction gives a brief overview of the field of synthetic biology and of societal reactions to it. The focus is on those lines of research that raise the most interesting ethical questions and that have (therefore) generated responses from academics, regulators and civil society. It introduces the four articles that constitute the main body of the dissertation, and provides a more in-depth discussion of objections concerning the human relationship to life and nature (which I only touch relatively briefly on in the articles).

Articles 1 and 2 discuss what I argue are shortcomings in the typical framing of the ethical questions that synthetic biology raises. In article 1, I argue that the creation of artificial life – which is one description of what synthetic biologist are doing – is morally significant. I do so in opposition to Douglas, Powell and Savulescu. I argue that their conceptions of ‘artificial life’ and ‘moral significance’ are unduly narrow, and lead to blindness to important objections. In article 2 I discuss arguments that defend synthetic biology and similar technologies on the basis of its similarity with established technologies, and show that there are ways to avoid them.

In article 3 and 4, I discuss aspects of risk regulation. Article 3 defends the precautionary principle from objections based on its supposed irrationality. Article 4 criticises Sunstein’s argument for cost-benefit analysis as a correction for cognitive biases. I argue that the goal of avoiding cognitive error does not warrant core aspects of cost-benefit analysis.

## References

- Abdellaoui, M. (2009). Rank-Dependent Utility. In P. Anand, P.K. Pattanaik & C. Puppe (Eds.), *The Handbook of Rational and Social Choice*. Oxford: Oxford University Press.
- Acevedo-Rocha, C.G. (2016). The Synthetic Nature of Biology. In K. Hagen, M. Engelhard & G. Toepfer (Eds.), *Ambivalences of Creating Life: Societal and Philosophical Dimensions of Synthetic Biology*. Cham: Springer.
- Alhakami, A.S. & Slovic, P. (1994). A Psychological Study of the Inverse Relationship Between Perceived Risk and Perceived Benefit, *Risk Analysis*, 14(6), 1085-1096.
- Ancilotti, M. & Eriksson, S. (2016). Synthetic Biology in the Press. In K. Hagen, M. Engelhard & G. Toepfer (Eds.), *Ambivalences of Creating Life: Societal and Philosophical Dimensions of Synthetic Biology*. Cham: Springer.
- Anderson, E. (1993). *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.
- Anderson, E. & Pildes, R. (2000). Expressive Theories of Law: A General Restatement. *University of Pennsylvania Law Review*, 148(5), 1503-1575.
- Arkes, H., Gigerenzer, G. & Hertwig, G. (2016). How Bad Is Incoherence? *Decision*, 3(1), 20-39.
- Attfield, R. (2012). Biocentrism and Artificial Life, *Environmental Values*, 21(1), 83-94.

Baertschi, B. (2012a). Defeating the Argument from Hubris. *Bioethics*, 27(8), 435-441.

Baertschi, B. (2012b). The Moral Status of Artificial Life, *Environmental Values*, 21(1), 5-18.

Bar-Yam, S., Byers-Corbin, J., Casagrande, R., Eichler, F., Lin, A., Oesterreicher, M. ... Oye, K.A. (2012). *The Regulation of Synthetic Biology: A Guide to United States and European Union Regulations, Rules and Guidelines*. Retrieved from [https://www.synberc.org/sites/default/files/Concise%20Guide%20to%20Synbio%20Regulation%20OYE%20Jan%202012\\_0.pdf](https://www.synberc.org/sites/default/files/Concise%20Guide%20to%20Synbio%20Regulation%20OYE%20Jan%202012_0.pdf)

Basl, J. & Sandler, R. (2013). The good of non-sentient entities: Organisms, artifacts and synthetic biology, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 697-705.

Bedau, M., Church, G., Rasmussen, S., Caplan, A., Benner, S., Fussenegger, M., Collins, J., & Deamer, D. (2010). Life after the Synthetic Cell. *Nature*, 465, 422-424.

Bedau, M., Parke, E., Tangen, U. & Hantsche-Tange, B. (2009). Social and ethical checkpoints for bottom-up synthetic biology, or protocells, *Systems and Synthetic Biology*, 3, 65-75.

Benner, S.A. & Sismour, A.M. (2005). Synthetic Biology, *Nature Reviews Genetics*, 6, 533-543.

- Berg, N., Eckel, C. & Johnson, C. (2010). *Inconsistency Pays? Time-inconsistent subjects and EU violators earn more* (MPRA Working Paper). Retrieved from <https://mpa.ub.uni-muenchen.de/26589/>
- Bergeson, L.L., Campbell, L.M., Dolan, S.L., Engler, R.E., Baron, K.F., Auerbach, B. ... Auer, C.M. (2015). *The DNA of the U.S. Regulatory System: Are We Getting It Right for Synthetic Biology*. Retrieved from [https://www.wilsoncenter.org/sites/default/files/synbio\\_reg\\_report\\_final.pdf](https://www.wilsoncenter.org/sites/default/files/synbio_reg_report_final.pdf)
- Blackburn, S. (1996). Securing the notes: Moral epistemology for the quasi-realist'. In W. Sinnott-Armstrong & M. Timmons (Eds.) *Moral Knowledge? New Readings in Moral Epistemology*. Oxford: Oxford University Press.
- Boldt, J. (2013a). Life as a Technological Product: Philosophical and Ethical Aspects of Synthetic Biology. *Biological Theory*, 8, 391-401
- Boldt, J. (2013b). Creating Life: Synthetic Biology and Ethics. In G.E. Kaebnick & T.H. Murray (Eds.), *Synthetic Biology and Morality: Artificial Life and the Bounds of Nature*. Cambridge, MA: MIT Press.
- Boldt, J. (2013c). Do We Have a Moral Obligation to Synthesize Organisms to Increase Biodiversity? On Kinship, Awe and the Value of Life's Diversity, *Bioethics*, 27(8), 411-418.
- Boldt, J., & Müller, O. (2008). Newtons of the Leaves of Grass. *Nature Biotechnology* 26, 387-389.

- Boyle, R.J. & Savulescu, J. (2001). Ethics of using preimplantation genetic diagnosis to select a stem cell donor for an existing person, *British Medical Journal*, 323, 1240-1243.
- Brighton, H. & Gigerenzer, G. (2012). Homo Heuristicus and the Bias-Variance Dilemma. In J. Schulkin (Ed.): *Action, Perception and the Brain: Adaptation and Cephalic Expression*. New York: Palgrave Macmillan.
- Brown, C. (2011). Two Kinds of Holism About Values, *The Philosophical Quarterly*, 57, 456-463.
- Bulleit, R.W. (2005). *Hunters, Herders and Hamburgers: The Past and Future of Human-Animal Relationships*. Columbia University Press, New York.
- Butler, D. (2014). Biosafety controls come under fire, *Nature*, 511, 515-516.
- Callicott, J.B. (1980). Animal Liberation: A Triangular Affair, *Environmental Ethics*, 2, 311-338.
- Caplan, A. (2009). Moving ahead but with greater controls, *Nature Biotechnology*, 27(12), 1105.
- Caplan, A. (2010, September 13). *Testimony to the Presidential Commission for the Study of Bio-ethical Issues*. Retrieved from <http://bioethics.gov/sites/default/files/Testimony-of-Arthur-L-Caplan.pdf>



- Cello, J., Paul, A.V. & Wimmer, E. (2002): Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template, *Science*, 297, 1016-1018.
- Cho, M., Magnus, D., Caplan, A.C., McGee, D. & the Ethics of Genomics Group (1999). Ethical Considerations in Synthesizing a Minimal Genome, *Science*, 286, 2087 + 2089-2090.
- Church, G.M. (2004, August 6). *A Synthetic Biohazard Non-proliferation Proposal*. Retrieved from [http://arep.med.harvard.edu/SBP/Church\\_Biohazard04c.htm](http://arep.med.harvard.edu/SBP/Church_Biohazard04c.htm)
- Cohen, G.A. (2011). Rescuing Conservatism. In R.J. Wallace, R. Kumar & S. Freeman (Eds.), *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. Oxford: Oxford University Press.
- Cohen, I.G., Daniels, N. & Eyal, N. (Eds.) (2015): *Identified Versus Statistical Lives: An Interdisciplinary Perspective*. Oxford: Oxford University Press.
- Comstock, G. (2010). Ethics and Genetically Modified Food. In F.T. Gottwald, H.W. Ingen-siep & M. Meinhard (Eds.) *Food Ethics* (pp. 49-66). Dordrecht: Springer.
- Dancy, J. (1993). *Moral Reasons*. Cambridge, MA: Blackwell.
- Deckers, J. (2005): Are Scientists Right and Non-Scientists Wrong? Reflections on Discussions of GM. *Journal of Agricultural and Environmental Ethics*, 18, 451-478

- DeLoache, W.C., Russ, Z.N., Narcross, L., Gonzales, A.M., Martin, V.J.J. & Dueber, J.E. (2015). An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose, *Nature Chemical Biology*, 11, 465-471.
- DeMiguel, V., Garlappi, L. & Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22(5), 1915-1953.
- Deplazes-Zemp, A. (2012). The moral impact of synthesising living organisms: Biocentric views on synthetic biology, *Environmental Values*, 21(1), 63-82.
- Douglas, T., Powell, R., & Savulescu, J. (2013). Is the Creation of Artificial Life Morally Significant? *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 688-696.
- Douglas, T. & Savulescu, J. (2010). Synthetic biology and the ethics of knowledge, *Journal of Medical Ethics*, 36, 687-693.
- Dormitzer, P.R., Suphaphiphat, P., Gibson, D.G., Wentworth, D.E., Stockwell, T.B., Algire, M.A. ... Venter, J.C. (2013). Synthetic Generation of Influenza Vaccine Viruses for Rapid Response to Pandemics, *Science Translational Medicine*, 5(185ra68), 1-12.
- Dörner, D. (1990). The Logic of Failure, *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 327, 463-473.
- Ehrenberg, R. (2015). Engineered yeast paves way for home-brew heroin, *Nature*, 521, 267-268.

Elliot, R. (1982). Faking Nature, *Inquiry*, 25, 81-93.

Endy, D. (2005). Foundations for Engineering Biology. *Nature*, 438, 449-453.

ETC Group (2014, June 2). *Open Letter to Ecover/Method*. Retrieved from <http://www.etcgroup.org/content/open-letter-ecover-method>

European Commission (2000, February 2). *Communication from the Commission on the Precautionary Principle*. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52000DC0001&from=EN>

European Commission (2014, September 23). *Opinion on Synthetic Biology I: Definition*. Retrieved from [http://ec.europa.eu/health/scientific\\_committees/emerging/docs/scenihr\\_o\\_044.pdf](http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_044.pdf)

European Commission (2015a, April 29). *Opinion on Synthetic Biology II: Risk assessment methodologies and safety aspects*. Retrieved from [http://ec.europa.eu/health/scientific\\_committees/emerging/docs/scenihr\\_o\\_048.pdf](http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_048.pdf)

European Commission (2015b, November 27). *Opinion on Synthetic Biology I: Risks to the environment and biodiversity*. Retrieved from [http://ec.europa.eu/health/scientific\\_committees/emerging/docs/scenihr\\_o\\_050.pdf](http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_050.pdf)

European Group on Ethics in Science and New Technology (2009, November 17). *Ethics of Synthetic Biology*. Retrieved from [http://ec.europa.eu/archives/bepa/european-group-ethics/docs/opinion25\\_en.pdf](http://ec.europa.eu/archives/bepa/european-group-ethics/docs/opinion25_en.pdf)

Finkel, A.M. (2014). The Cost of Nothing Trumps the Value of Everything: The Failure of Regulatory Economics to Keep Pace with Improvements in Quantitative Risk Analysis, *Michigan Journal of Environmental and Administrative Law*, 4(1), 91-156.

FOE, ICTA & the ETC Group (2012, March 13). *Principles for the Oversight of Synthetic Biology*. Retrieved from [http://libcloud.s3.amazonaws.com/93/ae/9/2287/2/Principles\\_for\\_the\\_oversight\\_of\\_synthetic\\_biology.pdf](http://libcloud.s3.amazonaws.com/93/ae/9/2287/2/Principles_for_the_oversight_of_synthetic_biology.pdf)

Fox, C.R., Erner, C. & Walters, D.J. (2015). Decision Under Risk: From the Field to the Laboratory and Back, in G. Keren & G. Wu (Eds.): *The Wiley Blackwell Handbook of Judgment and Decision Making*. Chichester: John Wiley & Sons.

French, C.E., de Mora, K., Joshi, N., Elfick, A., Haseloff, J. & Ajioka, J. (2011). Synthetic Biology and the Art of Biosensor Design. In E.R. Choffnes, D.A. Relman & L. Pray (Eds.), *The Science and Applications of Synthetic and Systems Biology*. Washington, DC: The National Academies Press.

Frick, J. (2015). Contractualism and Social Risk, *Philosophy and Public Affairs*, 43(3), 175-223.

Garfinkel, M.S., Endy, D., Epstein, G.L. & Friedman, R.M. (2007). *Synthetic Genomics: Options for Governance*. Retrieved from <http://www.jcvi.org/cms/fileadmin/site/research/projects/synthetic-genomics-report/synthetic-genomics-report.pdf>

Garnett, K. & Parsons, D.J. (2016). Multi-Case Review of the Application of the Precautionary Principle in European Union Law and Case Law, *Risk Analysis*, doi:10.1111/risa.12633

Gaskell, G., Stares, S., Allansdottir, A., Allum, N., Castro, P., Esmer, Y., ... Wagner, W. (2010). Europeans and biotechnology in 2010: Winds of change? *Report to the European Commissions Directorate-General for Research*. Retrieved from [http://ec.europa.eu/public\\_opinion/archives/ebs/ebs\\_341\\_winds\\_en.pdf](http://ec.europa.eu/public_opinion/archives/ebs/ebs_341_winds_en.pdf).

Gaus, G. (2005). The Place of Autonomy within Liberalism. In J. Christman & J. Anderson (Eds.), *Autonomy and the Challenges to Liberalism*. Cambridge: Cambridge University Press.

Gelanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F. & Smolke, C.D. (2015). Complete biosynthesis of opioids in yeast, *Science*, 349, 1095-1100.

Gibbard, A. (1986). Risk and Value. in D. MacLean (Ed.), *Values at Risk*. Savage, MD: Rowman & Littlefield.

Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.-Y., Algire, M.A. ... Venter, J.C. (2010). Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome, *Science*, 329, 52-56.

Gigerenzer, G. & Selten, R. (Eds.) (2001). *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.

Goodin, R. (1985). Nuclear Disarmament and Moral Certainty, *Ethics*, 95(3), 641-658.

Goodpaster, K. (1978). On Being Morally Considerable. *The Journal of Philosophy*, 75(6), 308-325

Harbo, T.-I. (2010). The Function of the Proportionality Principle in EU Law, *European Law Journal*, 16(2), 158-185.

Hart Research Associates (2013, March 6). *Awareness & Impressions Of Synthetic Biology*. Report to the Synthetic Biology Project of the Woodrow Wilson International Center for Scholars. Retrieved from <https://www.wilsoncenter.org/publication/awareness-impressions-synthetic-biology>.

Hart Research Associates (2014, April 18). *Perceptions of Synthetic Biology and Neural Engineering*. Report to the Synthetic Biology Project of the Woodrow Wilson International Center for Scholars. Retrieved from [http://www.synbioproject.org/site/assets/files/1293/focusgroup\\_2014.pdf](http://www.synbioproject.org/site/assets/files/1293/focusgroup_2014.pdf)

Heath, J. (2011). Three Normative Models of the Welfare State, *Public Reason*, 3(2), 13-43.

Heinzerling, L. (1998). Regulatory Costs of Mythic Proportions, *The Yale Law Journal*, 107(7), 1981-2070.

Heinzerling, L. (2014). Inside EPA: A Former Insider's Reflections on the Relationship Between the Obama EPA and the Obama White House, *Pace Environmental Law Review* 31(1) 325-369.

Heinzerling, L. (2015). Statistical Lives in Environmental Law. In I.G. Cohen, N. Daniels & N. Eyal (Eds.), *Identified Versus Statistical Lives: An Interdisciplinary Perspective*. Oxford: Oxford University Press.

Hertwig, R., Pachur, T. & Kurzenhäuser, S. (2005). Judgments of Risk Frequencies: Tests of Possible Cognitive Mechanisms, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(4), 621-642.

Heavey, P., (2013). Synthetic Biology Ethics: A Deontological Assessment, *Bioethics*, 27(8), 442-452.

Heyd, D. (2012). Is there anything unique in the ethics of synthetic biology? *Perspectives in Biology and Medicine*, 55(4), 55:581–589.

Holm, S. (2012a). The Scientific Aspirations of Synthetic Biology and the Need for Analytic Ethics. *Ethics, Policy & Environment*, 15(1), 25-28.

Holm, S. (2012b). Biological Interests, Normative Functions, and Synthetic Biology, *Philosophy and Technology*, 25, 525-541.

- Holtug, N. (2009). Creating and Patenting New Life Forms. In H. Kuhse, P. Singer P (Eds.), *A Companion to Bioethics* (2<sup>nd</sup> ed.) (pp. 235-244). Malden, MA: Wiley-Blackwell.
- Horty, J.F. (2007). Reasons as Defaults. *Philosophers' Imprint*, 7, 1–28.
- Horty, J.F. (2012) *Reasons as Defaults*. Oxford: Oxford University Press.
- Huesken, S. (2014). Artificial Life and Ethics, *Nanoethics*, 8,111-116
- Hutchison, C.A., Chuang, R.-Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H. ... Venter, J.C. (2016). Design and synthesis of a minimal bacterial genome, *Science*, 351, 1414.
- James, A. (2012). Contractualism's (Not So) Slippery Slope, *Legal Theory*, 18(3), 263-292.
- Jefferson, C., Lentzos, F. & Marris, C. (2014). Synthetic biology and biosecurity: challenging the myths, *Frontiers in Public Health*, 2, 1-15.
- Jenni, K.E. & Lowenstein, G. (1997). Explaining the "Identifiable Victim Effect", *Journal of Risk and Uncertainty*, 14, 235-257.
- Jermy, A. (2011). Synthetic biology: Licensing bacteria to kill, *Nature Reviews Microbiology*, 9, 699.



- Jiang, P. (2014). A Uniform Precautionary Principle Under EU Law, *PKU Transnational Law Review*, 2(2), 491-518.
- John, S. (2007). How to take deontological concerns seriously in risk–cost–benefit analysis: a re-interpretation of the precautionary principle, *Journal of Medical Ethics*, 33, 221-224.
- Johnson, N. (2014, June 5). The next front in the GMO war: Synthetic biology. *Grist*. Retrieved from <http://grist.org/business-technology/the-next-front-in-the-gmo-war-synthetic-biology/>
- Johnson, N.P.A.S. & Mueller, J. (2002). Updating the Accounts: Global Mortality of the 1918-1920 “Spanish” Influenza Pandemic, *Bulletin of the History of Medicine*, 76(1), 105-115.
- Kaebnick, G.E. (2009). Should moral objections to synthetic biology affect public policy? *Nature Biotechnology*, 27(12), 1106-1108.
- Kaebnick, G.E. (2012, March 2012). *Carefully Precautionary about Synthetic Biology*. [Web log post]. Retrieved from <http://www.thehastingscenter.org/carefully-precautionary-about-synthetic-biology/#ixzz33iPzJJ5c>
- Kaebnick, G.E. & Murray, T.H. (2013). Introduction. In G.E. Kaebnick & T.H. Murray (Eds.), *Synthetic Biology and Morality: Artificial Life and the Bounds of Nature*. Cambridge, MA: MIT Press.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Penguin Books.

- Kaiser, J. (2014, May 27). U.S. advisers sign of on plan for reviewing risky virus studies, *ScienceInsider*. Retrieved from <http://www.sciencemag.org/news/2016/05/us-advisers-sign-plan-reviewing-risky-virus-studies>
- Kamm, F.M. (2007). *Intricate Ethics: Rights, Responsibilities and Permissible Harm*. Oxford: Oxford University Press.
- Katz, E. (1992). The Big Lie: Human Restoration of Nature, *Research in Philosophy and Technology*, 12, 231-241.
- Keasling, J. (2010). Manufacturing Molecules Through Metabolic Engineering, *Science*, 330, 1355-1358.
- Khalil, A.S. & Collins, J.J. (2010). Synthetic biology: applications come of age, *Nature Reviews Genetics*, 11, 367-379.
- Kronberger, N. (2012): Synthetic biology: taking a look at a field in the making, *Public Understanding of Science*, 21, 130-133.
- Kysar, D.A. (2004). Preferences for Processes: The Process/Product Distinction and the Regulation of Consumer Choice, *Harvard Law Review*, 118(2), 525-642.

- 'Labels for GMO Foods Are a Bad Idea' (2013, Sep 1). [Editorial] *Scientific American*, retrieved from <http://www.scientificamerican.com/article/labels-for-gmo-foods-are-a-bad-idea/>.
- Landes, X. & Holtug, N. (2015). Insurance, Equality and the Welfare State: Political Philosophy and (of) Public Insurance, *Res Publica*, 21, 111-118.
- Lee, K. (1999): *The Natural and the Artefactual: The Implications of Deep Science and Deep Technology for Environmental Philosophy*. Lanham, MD: Lexington Books.
- Lenman, J. (2000). Contractualism and Risk Imposition, *Politics, Philosophy and Economics*, 7(1), 99-122.
- Lewens, T. (2013). From *bricolage* to BioBricks™: Synthetic biology and rational design, *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 641-648.
- Lewidow, L., Carr, S. & Wield, D. (2005). European regulation of agri-biotechnology: precautionary links between science, expertise and policy, *Science and Public Policy*, 32(4), 261-276.
- Lewontin, R.C. (2014, May 8). The New Synthetic Biology: Who Gains? *New York Review of Books*. Retrieved from <http://www.nybooks.com/articles/2014/05/08/new-synthetic-biology-who-gains/>

- Liechtenstein, S. Slovic, P., Fischhoff, B., Layman, M. & Combs, B. (1978). Judged Frequency of Lethal Events, *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 551-578.
- Link, H.-J. (2013): Playing God and the Intrinsic Value of Life: Moral Problems for Synthetic Biology? *Science and Engineering Ethics*, 19, 435-448.
- Lipstich, M. & Ingleby, T.V. (2015). Reply to “Studies on Influenza Virus Transmission Between Ferrets: the Public Health Risks Revisited”, *mBio*, 6(1), 1-6.
- Lo, Y.S. (2010). The Land Ethic and Callicott’s Ethical System (1980-2001): An Overview and Critique. *Inquiry: An Interdisciplinary Journal of Philosophy*, 44(3), 331-358.
- Löfstedt, R. (2014). The precautionary principle in the EU: Why a formal review is long overdue, *Risk Management*, 16(3), 137-163.
- Machery, E. (2012) Why I Stopped Worrying about the Definition of Life... and Why You Should as Well. *Synthese*, 185, 145-164.
- Manson, N.A. (2002). Formulating the Precautionary Principle, *Environmental Ethics*, 24, 263-274.
- Marris, C., Wynne, B., Simmons, P., Weldon, S. (2001). Public Perceptions of Agricultural Biotechnologies in Europe. *Final Report of the PABE research project*. Retrieved from [www.csec.lancs.ac.uk/archive/pabe/docs/pabe\\_finalreport.pdf](http://www.csec.lancs.ac.uk/archive/pabe/docs/pabe_finalreport.pdf).

- McGarity, T.O. (2002a). Seeds of Distrust: Federal Regulation of Genetically Modified Foods, *University of Michigan Law Review*, 35(3), 403-510.
- McGarity, T.O. (2002b). Professor Sunstein's Fuzzy Math, *The Georgetown Law Journal*, 90, 2341-2377.
- McGarity, T.O. (2013). EPA at Helm's Deep: Surviving the Fourth Attack on Environmental Law, *Fordham Environmental Law Review*, 24(2), 205-241.
- McKibben, B. (2003). *The End of Nature* (revised ed.). London: Random House.
- McMahan, J. (1986). Nuclear Deterrence and Future Generations. In A. Cohen & S. Lee (Eds.): *Nuclear Weapons and the Future of Humanity: The Fundamental Questions*. Totowa, NJ: Rowman & Allanheld.
- McMahan, J. (2002). *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.
- Mielby, H., Sandøe, P., Lassen, J. (2013). Multiple aspects of unnaturalness: Are cisgenic crops perceived as being more natural and more acceptable than transgenic crops? *Agriculture and Human Values*, 30, 471-480.
- Miller, S. & Selgelid, M.J. (2007). Ethical and Philosophical Consideration of the Dual-Use Dilemma in the Biological Sciences, *Science and Engineering Ethics*, 13, 523-580.

- Moussaïd, M., Brighton, H. & Gaissmaier, W. (2015). The Amplification of Risk in Experimental Diffusion Chains, *PNAS*, 112(18), 5631-5636.
- Munthe, C. (2011). *The Price of Precaution and the Ethics of Risk*. Dordrecht: Springer.
- Munthe, C. (2015). Precaution, bioethics and normative justification, *Monash Bioethics Review*, 33, 219-225.
- Murphy, D.J. (2007). *People, Plants and Genes: The Story of Crops and Humanity*. Oxford: Oxford University Press.
- Møller, B.L. (2014). Disruptive innovation: Channeling photosynthetic electron flow into light-driven synthesis of high-value products, *Synthetic Biology*, 1, 330-359.
- National Research Council (2004). *Biotechnology Research in an Age of Terrorism*. Retrieved from <http://www.nap.edu/catalog/10827/biotechnology-research-in-an-age-of-terrorism>
- NSABB (2012, March 29-30). *Findings and Recommendations*. Retrieved from [https://www.nih.gov/sites/default/files/about-nih/nih-director/statements/collins/03302012\\_NSABB\\_Recommendations.pdf](https://www.nih.gov/sites/default/files/about-nih/nih-director/statements/collins/03302012_NSABB_Recommendations.pdf)
- Nozick, R. (1974). *Anarchy, State and Utopia*. New York, NY: Basic Books.

Nuffield Council on Bioethics (2012, December). *Emerging biotechnologies: Technology, choice and the public good*. Retrieved from <http://nuffieldbioethics.org/project/emerging-biotechnologies/>

Office of Management and Budget (2012). *2012 Report to Congress on the Benefits and Costs of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities*. Retrieved from [https://www.whitehouse.gov/sites/default/files/omb/inforeg/2012\\_cb/2012\\_cost\\_benefit\\_report.pdf](https://www.whitehouse.gov/sites/default/files/omb/inforeg/2012_cb/2012_cost_benefit_report.pdf)

O'Malley, M., Powell, A., Davies, J.F., & Calvert, J. (2008) Knowledge-Making Distinctions in Synthetic Biology. *BioEssays*, 30(1), 57-65.

Osterholm, M.T. & Kelley, N.S. (2012). Mammalian-Transmissible H5N1 Influenza: Facts and Perspective, *mBio*, 3(2), 1-4.

Pamlin, D. & Armstrong, S. (2015). *12 Risks that threaten human civilization: The case for a new risk category*. Global Challenges Foundation. Retrieved from <http://globalchallenges.org/wp-content/uploads/12-Risks-with-infinite-impact.pdf>

Parens, E., Johnston, J. & Moses, J. (2008). Do We Need "Synthetic Bioethics"? *Science*, 321, 1449.

Parens, E., Johnston, J. & Moses, J. (2009, June). *Ethical Issues in Synthetic Biology*. Retrieved from <http://www.synbioproject.org/publications/synbio3/>

Parker, R.W. (2003). Grading the Government, *The University of Chicago Law Review*, 70(4), 1345-1486.

Pauwels, E. (2009): Review of quantitative and qualitative studies on U.S. public perceptions of synthetic biology. *Systems and Synthetic Biology*, 3, 37-46.

PCSBI (2010). *New Directions: The Ethics of Synthetic Biology and Emerging Technologies*. Washington, DC: Presidential Commission for the Study of Bioethical Issues.

Perkel, J.M. (2015, June 9). Revolutionizing Vaccine Development with Synthetic Biology. *Biocompare*. Retrieved from <http://www.biocompare.com/Editorial-Articles/174945-Revolutionizing-Vaccine-Development-with-Synthetic-Biology/>

Pierce, R.L. (2012). Whose ethics of knowledge? Taking the next step in evaluating knowledge in synthetic biology: a response to Douglas and Savulescu, *Journal of Medical Ethics*, 38, 636-638.

Peterson, M. (2003). Transformative Decision Rules, *Erkenntnis*, 58(1), 71-85.

Peterson, M. (2006). The Precautionary Principle is Incoherent, *Risk Analysis*, 26(3), 595-601.

Peterson, M. (2008). *Non-Bayesian Decision Theory: Beliefs and Desires as Reasons for Action*. Dordrecht: Springer.



Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge: Cambridge University Press.

Pollack, A. (2002, July 12). Scientists Create a Live Polio Virus. *New York Times*. Retrieved from <http://www.nytimes.com/2002/07/12/science/12POLI.html?pagewanted=all>

Pollack, A., (2013, October 20). What's That Smell? Exotic Scents Made From Re-engineered Yeast. *New York Times*. Retrieved from [http://www.nytimes.com/2013/10/21/business/whats-that-smell-exotic-scents-made-from-re-engineered-yeast.html?\\_r=1](http://www.nytimes.com/2013/10/21/business/whats-that-smell-exotic-scents-made-from-re-engineered-yeast.html?_r=1)

Powell, R. (2015). Adopting a Technological Stance Toward the Living World. Promises, Pitfalls and Perils. In S.O. Hansson (Ed.), *The Role of Technology in Science*. Dordrecht: Springer.

Preston, B. (2013). Synthetic Biology as Red Herring. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, 649-659.

Preston, C.J. (2008). Synthetic Biology: Drawing a Line in Darwin's Sand. *Environmental Values*, 17(1), 23-39

Preston, C.J. (2013a). Synthetic Bacteria, Natural Processes and Intrinsic Value. In G.E. Kaebnick & T.H. Murray (Eds.), *Synthetic Biology and Morality: Artificial Life and the Bounds of Nature*. Cambridge, MA: MIT Press.

Preston, C.J. (2013b). Evolution and the Deep Past: Intrinsic Responses to Synthetic Biology.

In R.L. Sandler (Ed.), *Ethics and Emerging Technologies*. Basingstoke: Palgrave Macmillan.

Pugh, J. (2015a, December 2). The Ethics of Genetically Modified Mosquitoes and Gene-

Drive Technology. [Web log post]. Retrieved from <http://blog.practicaethics.ox.ac.uk/2015/12/the-ethics-of-genetically-modified-mosquitoes-and-gene-drive-technology/>.

Pugh J (2015b, December 2). Playing God with mosquitoes? We humans have loftier aims.

*The Conversation*. Retrieved from <https://theconversation.com/playing-god-with-mosquitoes-we-humans-have-loftier-aims-51362>.

Randall, A. (2011). *Risk and Precaution*. Cambridge: Cambridge University Press.

Rappert, B. & Selgelid, M.J. (Eds.) (2013). On the Dual Uses of Science and Ethics: Principles, Practices and Prospects. Canberra: ANU Press.

Reardon, S. (2014). US suspends risky disease research, *Nature*, 514, 411-412.

Ribeiro, Silvia & Thomas, Jim (2015, October 23): A Bad Bet on Synthetic Biology. *Project Syndicate*. Retrieved from <https://www.project-syndicate.org/commentary/stevia-synthetic-biology-dangers-by-silvia-ribeiro-and-jim-thomas-2015-10>.

Rodemeyer, M. (2009). *New Life, Old Bottles: Regulating First-Generation Products of Synthetic Biology*. Retrieved from [http://www.synbioproject.org/process/assets/files/6319/nano\\_synbio2\\_electronic\\_final.pdf](http://www.synbioproject.org/process/assets/files/6319/nano_synbio2_electronic_final.pdf)

- Rozell, D.J. (2015). Assessing and Managing the Risks of Potential Pandemic Pathogen Research, *mBiom* 6(4), 1-4.
- Rozin, P. (2005). The Meaning of “Natural”: Process More Important Than Content. *Psychological Science*, 16(8), 652-658.
- Rozin, P., Spranca, M., Krieger, Z., Neuhaus, R., Surillo, D., Swerdlin, A. & Wood, K. (2004). Preference for natural: instrumental and ideational/moral motivations, and the contrast between foods and medicines. *Appetite*, 43, 147-154
- Rozin, P., Fischler, C., Shields-Argelès, C. (2009). Additivity Dominance: Additives are more potent and more often lexicalized across languages than “subtractives”. *Judgment and Decision Making*, 4(5), 475-478
- Sandin, P. (1999). Dimensions of the Precautionary Principle, *Human and Ecological Risk Assessment*, 5(5), 889-907.
- Sandin, P., Peterson, M., Hansson, S.O., Rudén, C. & Juthe, A. (2002). Five charges against the precautionary principle, *Journal of Risk Research*, 5(4), 287-299.
- Sandler, R. (2004). An Aretaic Objection to Agricultural Biotechnology, *Journal of Agricultural and Environmental Ethics*, 17, 301-317.
- Sandler, R. (2007). *Character and Environment*. New York: Columbia University Press.

- Schmidt, M., Ganguli-Mitra, A., Torgersen, H., Kelle, A., Deplazes, A. & Biller-Andorno, N. (2009). A priority paper for the societal and ethical aspects of synthetic biology, *Systems and Synthetic Biology*, 3, 3-7.
- Scott, D. (2011). The Technological Fix Criticisms and the Agricultural Biotechnology Debate, *Journal of Agricultural and Environmental Ethics*, 24, 207-226.
- Schroeder, M. (2011). Holism, Weight and Undercutting, *Noûs*, 45(2), 328-344.
- Selgelid, M.J. (2007). A Tale of Two Studies: Ethics, Bioterrorism and the Censorship of Science, *Hastings Center Report*, 37(3), 35-43.
- Selgelid, M.J. (2016). Gain-of-Function Research: Ethical Analysis, *Science and Engineering Ethics*, doi:10.1007/s11948-016-9810-1
- Service, R.F. (2015, May 18). Researchers closer to engineering yeast that make morphine, spurring worries. *ScienceInsider*. Retrieved from <http://www.sciencemag.org/news/2015/05/researchers-closer-engineering-yeast-make-morphine-spurring-worries>
- Shue, H. (2015). Uncertainty as the Reason for Action: Last Opportunity and Future Climate Disaster, *Global Justice*, 8(2), 86-103.
- Slomovic, S., Pardee, K. & Collins, J.J. (2015). Synthetic biology devices for in vitro and in vivo diagnostics, *PNAS*, 112(47), 14429-14435.

Slovic, P. (2000). *The Perception of Risk*. London: Earthscan.

Smith, K. (2013). Synthetic Biology: A Utilitarian Perspective, *Bioethics*, 27(8), 453-463.

Sperber, D. (2007). Seedless Grapes: Nature and Culture. In E. Margolis & S. Laurence (Eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford: Oxford University Press.

Steel, D. (2013). The Precautionary Principle and the Dilemma Objection, *Ethics, Policy and Environment*, 16(3), 321-340.

Steel, D. (2014). *Philosophy and the Precautionary Principle: Science, Evidence and Environmental Policy*. Cambridge: Cambridge University Press.

Steurer, W. (2016). "Some Kind of Genetic Engineering... Only One Step Further"—Public Perceptions of Synthetic Biology in Austria. In K. Hagen, M. Engelhard & G. Toepfer (Eds.), *Ambivalences of Creating Life: Societal and Philosophical Dimensions of Synthetic Biology*. Cham: Springer.

Stiglitz, J.E. & Greenwald, B.N. (2014): *Creating a Learning Society*. New York: Columbia University Press.

Strawson, P.F. (1962). Freedom and Resentment, *Proceedings of the British Academy*, 48, 1-25.

- Sunstein, C.R. (2001). Cognition and Cost-Benefit Analysis, *The Journal of Legal Studies*, 29(52), 1059-1103.
- Sunstein, C.R. (2002). *Risk and Reason: Safety, Law and the Environment*. Cambridge: Cambridge University Press.
- Sunstein, C.R. (2005). *Laws of Fear: Beyond the Precautionary Principle*. Cambridge: Cambridge University Press.
- Sunstein, C.R. (2013). *Simpler: The Future of Government*. New York: Simon & Schuster.
- Sunstein, C.R. (2014). *Valuing Life: Humanizing the Regulatory State*. Chicago: The University of Chicago Press.
- Sunstein, C.R., Kahneman, D., Schkade, D. & Ritov, I. (2002). Predictably Incoherent Judgments, *Stanford Law Review*, 54(7), 1153-1215.
- Suter, R.S., Pachur, T. & Hertwig, R. (2015). How Affect Shapes Risky Choice: Distorted Probability Weighing Versus Probability Neglect, *Journal of Behavioral Decision Making*, doi: 10.1002/bdm.1888
- Tengs, T.O. & Graham, J.D. (1996). The Opportunity Cost of Haphazard Social Investments in Life-Saving. In R.W. Hahn (Ed.), *Risks, Costs and Lives Saved: Getting Better Results from Regulation*. Oxford: Oxford University Press.

- Thomas, J. (2013, April 12): Synthetic anti-malarial compound is bad news for artemisia farmers. *The Guardian*. Retrieved from <http://www.theguardian.com/global-development/poverty-matters/2013/apr/12/synthetic-malaria-compound-artemisia-farmers>
- Thompson, J. (2010). The Apology Paradox. *The Philosophical Quarterly*, 50, 470-475.
- Thompson, P.B. (2003). Unnatural Farming and the Debate over Genetic Manipulation. In V.V. Gehring (Ed.). *Genetic Prospects: Essays on Biotechnology, Ethics and Public Policy*. Lanham, MD: Rowman and Littlefield.
- Thompson, P.B. (2012). Synthetic Biology Needs A Synthetic Bioethics, *Ethics, Policy and Environment*, 15(1), 1-20.
- Thompson, P.B. (2015). From Synthetic Bioethics to One Bioethics: A Reply to Critics, *Ethics, Policy and Environment*, 18(2), 215-224.
- Thompson, R.P. (2011). *Agro-technology: A Philosophical Introduction*. Cambridge: Cambridge University Press.
- Thomson, J.J. (1986). *Rights, Restitution and Risk: Essays in Moral Theory*. Cambridge, MA: Harvard University Press.
- Trouwborst, A. (2006). *Precautionary Rights and Duties of States*. Leiden: Martinus Nijhoff.

Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solórzano, A., Swayne, D.E. ... Garcia-Sastre, A. (2005). Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus, *Science*, 310, 77-80.

UNEP (1992). *Rio Declaration on Environment and Development*. Retrieved from <http://www.unep.org/documents.multilingual/default.asp?documentid=78&articleid=1163>

Von Bubnoff, A. (2005). The 1918 flu virus is resurrected, *Nature*, 437, 794-795.

Watson, E. (2014, August 29). So Häagen-Dazs 'says no to synbio'? *Food Navigator*. Retrieved from <http://www.foodnavigator-usa.com/Manufacturers/So-Haeagen-Dazs-says-no-to-synbio>

Whaley, B.B. (1998) Evaluations of Rebuttal Analogy Users: Ethical and Competence Considerations. *Argumentation*, 12, 351-365

Wimmer, E., Mueller, S., Tumpey, T.M. & Taubenberger, J.K. (2009). Synthetic viruses: A new opportunity to understand and prevent viral disease, *Nature Biotechnology*, 27, 1163-1172.

Yang, S. (2015, May 18). Discovery paves way for homebrewed drugs, prompts call for regulation. *Berkeley News*. Retrieved from <http://news.berkeley.edu/2015/05/18/brewing-opioids-scientists-urge-regulation/>



